Author’s Note

The two Thesis –which are the subject matter of this document– are listed below. However, it is necessary to clarify that both Papers has been rewritten, broadening their scopes. Therefore a new section called *Proposal* has been added to both Thesis, in which I have tried to thoroughly assess both topics. I have also provided my opinion, from a critical point of view, using my acquired experience over the last few years.

The first Thesis is called *International Overview on the Legal Framework for Highly Automated Vehicles*, followed by the second one named *State of the Art of Extraction of Traffic Messages from Social Media*.

Technische Universität München - Lehrstuhl für Verkehrstechnik
Univ.-Prof. Dr.-Ing. Fritz Busch
Arcisstraße 21 80333 München, www.vt.bgu.tum.de

BACHELOR'S THESIS

International Overview on the Legal Framework for Highly Automated Vehicles

Author:

Diego Pacho Toubes

Supervision:

Dipl.-Ing. Martin Margreiter (TUM)

Dr.-Ing. Miquel Estrada (UPC)

Date of Submission: 2015-10-12

Technische Universität München - Lehrstuhl für Verkehrstechnik
Univ.-Prof. Dr.-Ing. Fritz Busch
Arcisstraße 21 80333 München, www.vt.bgu.tum.de

BACHELOR'S THESIS

of Diego Pacho Toubes

Date of Issue: 2015-05-12

Date of Submission: 2015-10-12

Topic: **International Overview on the Legal Framework for Highly Automated Vehicles**

Worldwide leading car makers and technology companies are working on the development of autonomous or highly automated vehicles (HAV). Several demonstrations in the past already showed the technical feasibility of such approaches. Some of these vehicles are nowadays already allowed to use (partially) the public road networks.

The focus of this Bachelor's Thesis is to give an overview on the legal conditions and the legal frameworks which exist in the different parts of the world for such HAV. Therefore an international literature review has to be conducted to determine the current problems and legal issues which appear in the field of autonomous vehicles or HAV as well as the current or planned legislation on that topic as well as proposed solutions.

The student will present intermediate results to the supervisor (Dipl.-Ing. Martin Margreiter) in the fifth, tenth and 15th week.

The student must hold a 20 minute presentation with discussion at the most two months after the submission of the thesis. The presentation will be considered in the final grade in cases where the thesis itself cannot be clearly evaluated.

Univ.-Prof. Dr.-Ing. Fritz Busch

Abstract

The evolution of Autonomous and automated technologies during the last decades has been constant and maintained. All of us can remember an old film, in which they shown us a driverless car, and we thought it was just an unreal object born of filmmakers imagination. However, nowadays Highly Automated Vehicles are a reality, even not in our daily lives. Hardly a day we don’t have news about Tesla launching a new model or Google showing the new features of their autonomous car. But don’t have to travel far away from our borders. Here in Europe we also can find different companies trying, with more or less success depending on with, not to be lagged behind in this race.

But today their biggest problem is not only the liability of their innovative technology, but also the legal framework for Highly Automated Vehicles. As a quick summary, in only a few countries they have testing licenses, which not allow them to freely drive, and to the contrary most nearly ban their use. The next milestone in autonomous driving is to build and homogeneous, safe and global legal framework.

With this in mind, this paper presents an international overview on the legal framework for Highly Automated Vehicles. We also present de different issues that such technologies have to face to and which they have to overcome in the next years to be a real and daily technology.

Table of Contents

1 Introduction

During the last years, both autonomous technologies and vehicles have grown exponentially due to resources and capital dedicate only and exclusively for such purpose. This innovative technology is conceived to transform our current way-of-life and the set of future's cities. Even Autonomous already exist and have been used during the last decade in some industries (e.g., storage, assembly line), the average citizen does not know –or is not used to– this innovative technology. For this reason, the use of autonomous vehicles is thought up to be one of the most important threshold of such technology, generating large benefits to the society.

As of today, and even several Autonomous trials have been performed in different counties, they have allowed to demonstrate the ability of such technology to improve road safety and reduce pollution. Due to the *robot-car* allows the vehicle to know the position of other cars and external agents (such motorcyclist, bicyclist or pedestrians) in real traffic, it controls the speed of the vehicle (acceleration or braking where applicable) reducing its sudden changes, reducing fuel consumption and consequently reducing greenhouse gas emissions. In addition, and due to a high percentage of traffic accidents are caused by human drivers, the use of Autonomous will allow decreasing this percentage and the number of fatalities. Other benefits will be expansion in mobility, reduction in traffic congestion (as a result of crashes) and the birth of future industries and opportunities for the whole society.

Despite all these advantages, the velocity of development and implementation of autonomous vehicles has suffered several changes in pace, due to the restrictions in road legislation that do not allow vehicles without a driver. The fact that *every single vehicle or convination or vehicles should have a driver* was stated for both the Geneva and the Vienna Convention on Road Traffic, treaties that are considered by many countries as the basis of road legislation. Thus, it is needed to review the international legal framework for Highly Automated Vehicles. In addition, and due to legislation is the main but not the only issue of such technology, another review is presented to identify other important problems that could arise when using autonomous vehicles. Finally, a guideline is presented to define and establish the basis of future "Autonomous-friendly" legislation and overcome all the outstanding challenges.

2 Legislative Background

If we want to know the first enacted laws regarding to the use of Autonomous, we have to move to two specifics dates. The first one was at the late 1949 in Geneva and the second one some years later in Vienna, specifically in 1968. Both treaties are considered the beginning of autonomous legislation on one side, and on the other side the main problem which some countries have to face up to. They contain laws which set different rules between vehicle, driver and their relationship in order to establish certain uniform rules. Nevertheless, they were intended more to the field of agriculture and stockbreeding than to autonomous. Also noteworthy is that both treaties do not categorically ban the use of such technologies otherwise they are only an obstacle to their legal implementation. Below are shown the features of both treaties and the different countries that are party –or not– of them.

2.1 Convention on Road Traffic, Geneva

According to the [1]note of the United Nations, the Convention was prepared and opened for signature by the United Nations Conference on Road and Motor Transport held at Geneva from 23 August to 19 September 1949. It promotes the development and safety of international road traffic by establishing certain uniform rules. It was ratified by 95 States, which are shown in Fig.2.1. It has eight chapters and ten annexes, but the part concerning the automated vehicles is found at *"Chapter II. Rules of the Road; Article 8*, which reads as follows:

> *1. Every vehicle or combination of vehicles proceeding as a unit shall have a driver.*
>
> *2. Draught, pack or saddle animals shall have a driver, and cattle shall be accompanied, except in special areas which shall be marked at the points of entry.*
>
> *3. Convoys of vehicles and animals shall have the number of drivers prescribed by domestic regulations.*
>
> *4. Convoys shall, if necessary, be divided into sections of moderate length, and be sufficiently spaced out for the convenience of traffic. This provision does not apply to regions where migration of nomads occurs.*

---

[1] *https://treaties.un.org/Pages/ViewDetailsV.aspx?src=TREATY&mtdsg_no=XI-B-1&chapter=11&Temp=mtdsg5&lang=en*

5. Drivers shall at all times be able to control their vehicles or guide their animals. When approaching other road users, they shall take such precautions as may be required for the safety of the latter.

In addition, we find relevant information at Article 4 that concerns Autonomous. It defines driver as "*any person who drives a vehicle, including cycles, or guides draught, pack or saddle animals or herds of flocks on a road, or who is in actual physical control of the same"*. It also defines motor vehicle as *"any self-propelled vehicle normally used for the transport of persons or goods upon a road, other than vehicles running on rails or connected to electric conductors"*, nevertheless vehicle is not defined.

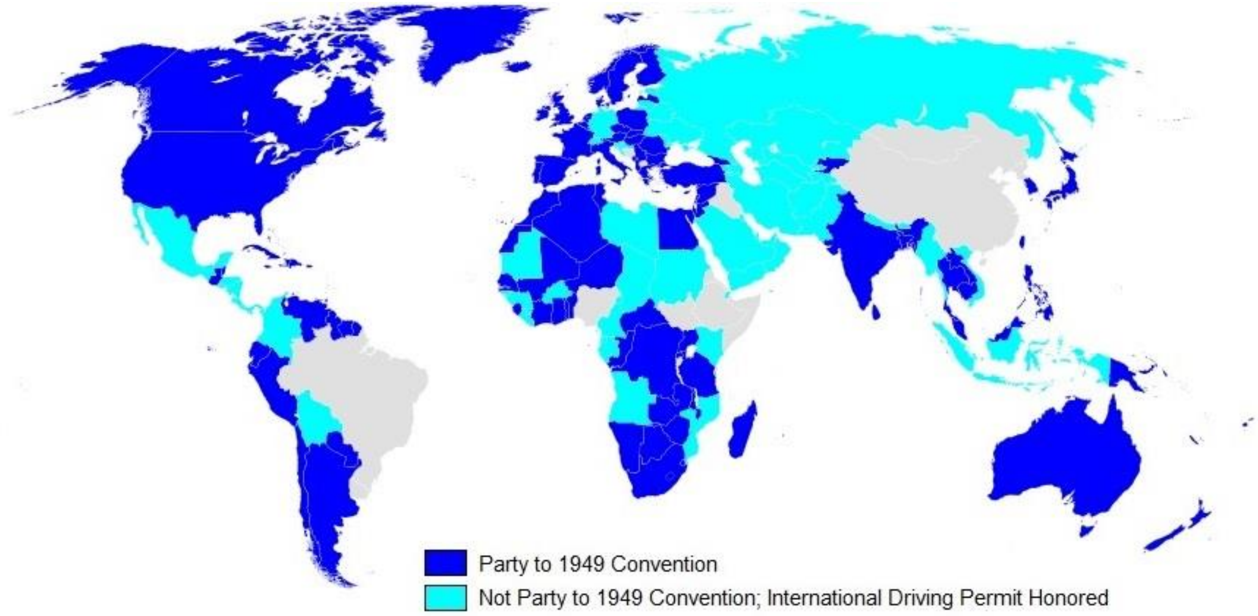

Fig.2.1 States which ratified the Geneva Convention

Source: Joe DeRose and Diego Pacho

2.2 Convention on Road Traffic, Vienna

Along with the [2]note of the United Nations, the Convention was prepared and opened for signature by the United Nations Conference on Road Traffic, held at Vienna from 7 October to 8 November 1968. It is an international treaty designed to facilitate international road traffic and to increase road safety by establishing standard traffic rules among the contracting parties. Until today, the convention has been ratified by 73 countries, shown in Fig.2.2. It

[2] *https://treaties.un.org/Pages/ViewDetailsIII.aspx?src=TREATY&mtdsg_no=XI-B-19&chapter=11&Temp=mtdsg3&lang=en*

consists in six chapters and seven annexes, but the part concerning the automated vehicles is found at *“Chapter II. Rules of the Road; Article 8. Drivers”* which reads as follows:

> *1. Every moving vehicle or combination of vehicles shall have a driver.*
>
> *2. It is recommended that domestic legislation should provide that pack, draught or saddle animals, and, except in such special areas as may be marked at the entry, cattle, singly or in herds, or flocks, shall have a driver.*
>
> *3. Every driver shall possess the necessary physical and mental ability and be in a fit physical and mental condition to drive.*
>
> *4. Every driver of a power-driven vehicle shall possess the knowledge and skill necessary for driving the vehicle; however, this requirement shall not be a bar to driving practice by learner-drivers in conformity with domestic legislation.*
>
> *5. Every driver shall at all times be able to control his vehicle or to guide his animals.*
>
> *6. A driver of a vehicle shall at all times minimize any activity other than driving. Domestic legislation should lay down rules on the use of phones by drivers of vehicles. In any case, legislation shall prohibit the use by a driver of a motor vehicle or moped of a hand-held phone while the vehicle is in motion.*

The sixth paragraph was added to this article in 2006. As before, there are two more articles which affect Autonomous. Article 1 defines driver as *“any person who drives a motor vehicle or other vehicle (including a cycle), or who guides cattle, singly or in herds, or flocks, or draught, pack or saddle animals on a road”*. Moreover Article 13 provides in part that *“every driver of a vehicle shall in all circumstances have his vehicle under control so as to be able to exercise due and proper care and to be at all times in a position to perform all manoeuvres required of him”*.

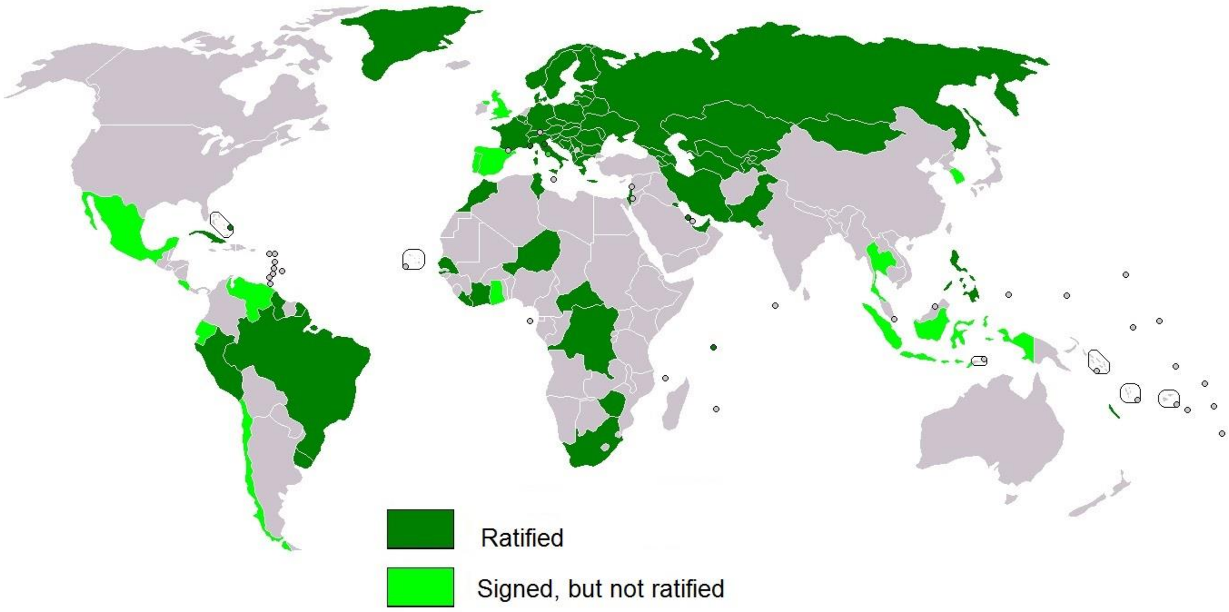

Fig.2.2 States which ratified the Vienna Convention

Source: Alinator (Wikipedia user) and Diego Pacho

3 International Overview

Several countries are promoting the development of autonomous technology by allowing driverless car into their roads, but in many cases with a lot of limitations. Almost all the trials have been performed under temporary license and through a given track. However, several public authorities of different countries have presented action plans with intent to encourage car manufacturers to carry on with autonomous technology development. At the same time, many announcements and demonstrations from automotive companies and research groups showed that the industry is globally moving closer to a scenario where the driving task will be gradually transferred from the human to the vehicle's smart systems.

As a quick overview, the first difference among the countries and their capacity to enact permissive Autonomous laws is their involvement in the Conventions of Vienna and Geneva. As explained above, both conventions established the first regulations framework of driverless cars, and as a consequence, signatory countries are ballasted by them. Also is important to underline that they were not thinking of Autonomous, but herd of animals. For this reason, some countries have taken them to be a barrier to the introduction of automated vehicles. Nowadays, both Conventions are in the process of being amended to allow a car to drive itself so long as the system can be overridden or switched off by the driver, though it has been argued that a further change is needed to allow automated vehicles on the roads in many countries.

Many countries are active in the field of autonomous vehicles, but the level and intensity of their activities vary. The scope of this section is to provide the reader an international overview of the different regulatory frameworks, depending on the country, and the fitting of autonomous initiatives in them.

3.1 United States of America

The United States of America is a federal republic, and therefore every State has their legal powers transferred. As a consequence, every State is responsible to enact new laws allowing or not the use of Autonomous on their roads. This has several consequences and one is the heterogeneity produced. At this point, the legal framework should be studied separately depending on the State. It is important to distinguish between States with autonomous concerning enacted legislation and only introduced or pending. The enacted legislation is shown at the Tab.3.1 with a short description of its content and other relevant information.

The information is obtained from two[3 and 4] different websites, which perfectly compile the already enacted laws in the United States of America. For this reason, I take this opportunity to make special mention of the valuable help that these two information sources have provide me and I want to express my gratitude for their extensive research in this topic.

ENACTED AUTONOMOUS VEHICLES LEGISLATION			
State	Bill Number	Description	Date
California	CA SB 1298	[5]Defines Autonomous Technology, Autonomous Vehicle and Operator. Requires the Department of the California Highway Patrol to adopt safety standards and performance requirements to ensure the safe operation and testing of Autonomous Vehicles, as defined, on the public roads in this state. Permits Autonomous Vehicles to be operated or tested on the public roads in this state pending the adoption of safety standards and performance requirements that would be adopted under this bill. Requires that the manufacturer of the autonomous technology installed on a vehicle shall provide a written disclosure to the purchaser of an autonomous vehicle, which describes what information is collected by the autonomous technology equipped on the vehicle.	Enacted and chaptered on September 25, 2012.

[3] *http://www.ncsl.org/research/transportation/autonomous-vehicles-legislation.aspx*

[4] *http://cyberlaw.stanford.edu/wiki/index.php/Automated_Driving:_Legislative_and_Regulatory_Action*

[5] *https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201120120SB1298*

District of Columbia	DC B19-0931	[6]Defines Autonomous Vehicle as a vehicle capable of navigating District roadways and interpreting traffic-control devices without a driver actively operating any of the vehicle's control systems. Requires a human driver prepared to take control of the autonomous vehicle at any moment and restricts conversion to recent vehicles and addresses liability of the original manufacturer of a converted vehicle. Final version removed previous provisions requiring autonomous vehicles to operate on alternative fuels and imposing a vehicle-miles-traveled tax in lieu of DoC motor fuel tax.	Enacted and chaptered on January 23, 2013.
Florida	FL HB 1207	[7]Defines Autonomous Vehicle and Autonomous Technology. Declares legislative intent to encourage the safe development, testing and operation of motor vehicles with autonomous technology on public roads of the state and finds that the state does not prohibit or specifically regulate the testing or operation of autonomous technology in motor vehicles on public roads. Authorizes a person who possesses a valid driver's license to operate an autonomous vehicle, specifying that the person who causes the vehicle's autonomous technology to engage is the operator. Authorizes the operation of autonomous vehicles by certain persons for testing purposes under	Enacted and chaptered on April 16, 2012.

[6] *http://dcclims1.dccouncil.us/lims/legislation.aspx?LegNo=B19-0931*

[7] *https://www.flsenate.gov/Session/Bill/2012/1207*

		certain conditions and requires an instrument of insurance, surety bond or self-insurance prior to the testing of a vehicle. Directs the Department of Highway Safety and Motor Vehicles to prepare a report recommending additional legislative or regulatory action that may be required for the safe testing and operation of vehicles equipped with autonomous technology, to be submitted no later than Feb. 12, 2014.	
	FL HB 599	[8]The relevant portions of this bill are identical to the substitute version of HB 1207.	Enacted and chaptered on April 29, 2012.
	FL SB 52	[9]Bans texting but exempts operators of autonomous vehicles operating in autonomous mode.	Enacted and chaptered on May 29, 2013.
Michigan	MI SB 0169	[10]Defines Automated Technology, Automated Vehicle and Automated Mode. Expressly permits testing of automated vehicles by certain parties under certain conditions. Defines Operator, addresses liability of the original manufacturer of a vehicle on which a third party has installed an automated system, directs state Department of Transportation with Secretary of State to submit report by Feb. 1, 2016.	Enacted and chaptered on December 20, 2013.

[8] *https://www.flsenate.gov/Session/Bill/2012/0599*

[9] *https://www.flsenate.gov/Session/Bill/2013/0052*

[10] *http://www.legislature.mi.gov/%28S%28h2gyj51e4ahtwx21ikmjdi0z%29%29/mileg.aspx?page=getobject&objectname=2013-SB-0169*

	MI SB 0663	[11]Declares that the original manufacturer of a vehicle is not liable for damages resulting from another person's conversion or attempted conversion of the vehicle into an automated motor vehicle, or the modification of installed equipment, unless the defect from which the damages resulted was present in the vehicle when it was manufactured. Similarly addresses liability of subcomponent system producers for equipment installed by those producers to convert vehicles into automated motor vehicles. Relates to automated mode conversions.	Enacted and chaptered on December 26, 2013.
Nevada	NV AB 511	[12]Authorizes operation of autonomous vehicles and a driver's license endorsement for operators of autonomous vehicles. Defines Autonomous Vehicle and directs state Department of Motor Vehicles (DMV) to adopt rules for license endorsement and for operation, including insurance, safety standards and testing.	Enacted and chaptered on June 17, 2011.
	NV SB 140	[13]Prohibits the use of cell phones or other handheld wireless communications devices while driving in certain circumstances, and makes it a crime to text or read data on a cellular phone while driving. Permits use of such devices for persons in a	Enacted and chaptered on June 17, 2011.

[11] *http://www.legislature.mi.gov/%28S%28jqj5cob0o5n0et413jyuwnvy%29%29/mileg.aspx?page=GetObject&objectname=2013-SB-0663*

[12] *https://www.leg.state.nv.us/Session/76th2011/reports/history.cfm?ID=1011*

[13] *http://leg.state.nv.us/76th2011/Reports/history.cfm?ID=324*

		legally operating autonomous vehicle. These persons are deemed not to be operating a motor vehicle for the purposes of this law.	
	NV SB 313	[14]Requires an autonomous vehicle that is being tested on a highway to meet certain conditions relating to a human operator. Requires proof of insurance. Prohibits an autonomous vehicle from being registered in the state, or tested or operated on a highway within the state, unless it meets certain conditions. Provides that the manufacturer of a vehicle that has been converted to be an autonomous vehicle by a third party is immune from liability for certain injuries. Adds exception in case of emergency to the statement that a person is not required to actively drive an autonomous vehicle.	Enacted and chaptered on June 2, 2013.
North Dakota	ND HB 1065	[15]Defines Autonomous Technology and Autonomous Vehicle. Establishes the requirements for insurance, testing and registration of autonomous vehicles. Delimits the liability of manufacturers for certain damages. Establishes regulations to authorize the operation of autonomous vehicles and encourages it. Also considers if current laws need to be changed to accommodate the introduction or testing of automated motor vehicles in North Dakota and any automated corridors affecting North Dakota.	Enacted and chaptered on March 20, 2015.

[14] *http://leg.state.nv.us/session/77th2013/reports/history.cfm?ID=759*

[15] *https://legiscan.com/ND/bill/1065/2015*

Tennessee	TN SB 598 TN HB 616	[16]Prohibits any political subdivision of the state from prohibiting within the jurisdictional boundaries of the political subdivision the use of a motor vehicle equipped with autonomous technology if the motor vehicle otherwise complies with all safety regulations of the political subdivision. Defines Autonomous Technology as technology that has the capability to drive a motor vehicle without the active physical control or monitoring by a human operator.	Enacted and chaptered on April 24, 2015.

Tab.3.1 Enacted autonomous legislation according to State

Source: Anne Teigen, Amanda Essex, Gabriel Weiner and Bryant Walker Smith

As a summary, and until today, only seven States and the District of Columbia have enacted laws which allow -in part- the use of Autonomous into their roads. Mostly define autonomous terminology, to take them into account in the legislation, and the rights and obligations of the parties concerned. The Fig.3.1 is useless to show the content explained above.

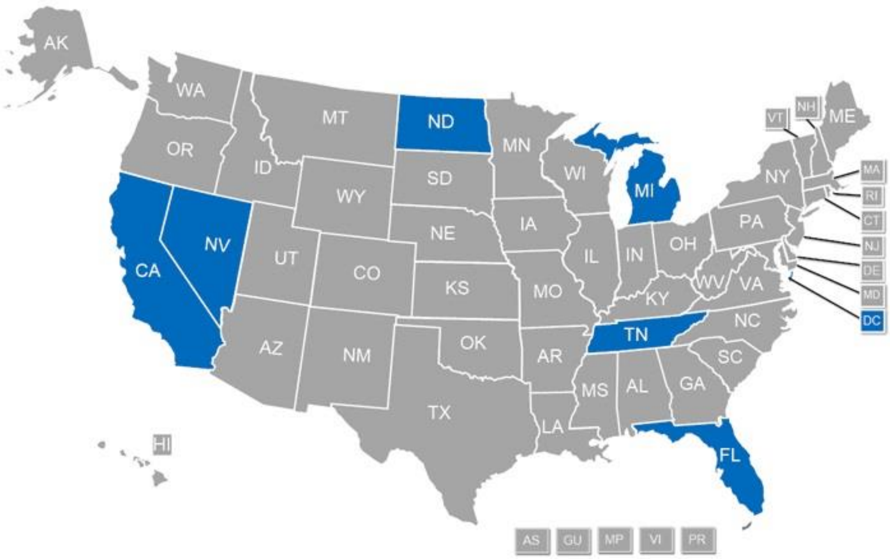

Fig.3.1 States with enacted autonomous legislation

Source: Anne Teigen and Amanda Essex

[16] *http://openstates.org/tn/bills/109/SB598/*

But not having enacted legislation does not mean that the State is turning it back to Autonomous. Several countries have introduced different laws concerning the use of Autonomous inside their jurisdictions, with more or less success. From 2011, when Nevada enacted the first law, different States have introduced new laws regarding Autonomous, being more and more every year. This has to be taken into consideration because of its velocity. In very little time one State can become a referent and another, which had already been, lag behind. Due to it velocity, new laws are enacted every day and that is why a continuous legislation review is needed. Next is shown a timeline of all the introduced legislation as of today. This table is extracted from the [17]article *"Autonomous | Self-Driving Vehicles Legislation"* to show how many countries have been trying to enact new legislations to regulate the use of autonomous vehicles.

INTRODUCED AUTONOMOUS VEHICLE LEGISLATION			
State	Bill Number	Description	Status
As of today, June 1st, 2015			
California	CA AB 1164	[18]Amends existing law addressing autonomous vehicles.	Pending –Passed House, Second Reading
	CA SB 431	[19]Amends existing law on following distance to allow for driver-assistive truck platooning.	Pending –Passed Senate
Connecticut	CT HB 6344	[20]Allows the use of autonomous vehicles for testing purposes and directs Department of Motor Vehicles to promulgate regulations concerning autonomous vehicles.	Failed - Died

[17] *http://www.ncsl.org/research/transportation/autonomous-vehicles-legislation.aspx*

[18] *https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201520160AB1164*

[19] *http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160SB431*

[20] *http://openstates.org/ct/bills/2015/HB6344/*

Georgia	GA SB 113	[21]Creates a new class of motor vehicles to be known as autonomous vehicles. Provides for definitions. Provides for requirements to operate an autonomous vehicle. Provides for the operation of autonomous vehicles on public highways for testing purposes. Provides for indemnity to vehicle manufacturers in certain instances. Provides for the regulation of autonomous vehicles. Provides for a penalty. Provides for related matters. Repeals conflicting laws.	Pending – Carryover
Hawaii	HI HB 632	[22]Authorizes for testing purposes the operation of autonomous vehicles in the state of Hawaii. Requires Department of Transportation to establish an application and approval process and report annually to the Legislature.	Failed First Crossover Deadline
	HI HB 1458	[23]Allows autonomous motor vehicles to be operated on any road, street, or highway if certain requirements are met. Allows for manufacturer	Failed First Crossover Deadline

[21] *http://www.legis.ga.gov/Legislation/en-US/display/20152016/SB/113*

[22] *http://openstates.org/hi/bills/2016%20Regular%20Session/HB632/*

[23] *http://2016-state-uas-legislation.silk.co/page/HI-H.B.-1458*

		testing of autonomous motor vehicles on any road, street or highway.	
	HI SB 630	[24]Allows a person who possesses a valid Hawaii driver license to operate an autonomous motor vehicle that employs autonomous technology. Defines Autonomous Vehicle and Autonomous Technology. Requires certain safety features. Specifies certain conditions for safety testing.	Failed First Crossover Deadline
Idaho	ID SB 1108	[25]Provides insurance requirements and requirements for testing autonomous driven vehicles. Provides that autonomous driven vehicles shall meet federal standards and regulations. Provides requirements for autonomous driven vehicles before testing or operation on highways within the state. Provides an exemption from liability for manufacturers and dealers and finally provides that all highways shall be open for testing.	Failed - Adjourned

[24] *http://www.capitol.hawaii.gov/measure_indiv.aspx?billtype=SB&billnumber=630*

[25] *https://legislature.idaho.gov/legislation/2015/S1108E1.pdf*

Illinois	IL HB 3136	[26]Amends the Vehicle Code and creates a new chapter on autonomous vehicles. Defines the terms, Autonomous Technology, Autonomous Vehicle, Secretary, Operator and Manufacturer. Provides for the operation of autonomous vehicles on public roads for testing purposes. Provides that an autonomous vehicle shall not be driven for other than testing purposes unless the manufacturer submits an application meeting certain requirements, and that application is approved by the secretary of state.	Pending –Passed House, Senate Third Reading
Maryland	MD HB 172	[27]Establishes the Task Force to Study Issues Related to the Use of Self Driving Vehicles. Provides for the composition, chair and staffing of the Task Force. Requires the Task Force to make specified determinations, review specified information, consider specified issues, and make specified recommendations related to the use of self driving vehicles.	Pending –Passed House, in Senate Judicial Proceedings Committee

[26] *http://www.ilga.gov/legislation/BillStatus.asp?DocTypeID=HB&DocNum=3136&GAID=13&SessionID=88&LegID=89280*

[27] *https://legiscan.com/MD/bill/HB172/2015*

	MD SB 778	[28]Requires the Task Force to report its findings and recommendations to the governor and the General Assembly on or before January 1, 2017.	Failed - Died
Massachusetts	MA HB 2977	[29]Authorizes the operation of autonomous vehicles without active control or monitoring by a human operator.	Pending –In Joint Committee on Transportation
	MA SB 1841	[30]Defines Autonomous Vehicles. Permits autonomous vehicles on public roads if they meet certain requirements. Permits operating the vehicle for testing. Requires the department of transportation to adopt regulations.	Pending –In Joint Committee on Transportation
Mississippi	MS SB 2676	[31]Authorizes the use of an autonomous motor vehicle in this state. Defines Autonomous Technology and Autonomous Vehicles. Provides for the safety and control of an autonomous vehicle. Provides for endorsement on driver's	Failed– Died in Committee

[28] https://mgaleg.maryland.gov/2015RS/fnotes/bil_0008/sb0778.pdf

[29] *https://malegislature.gov/Bills/189/House/H2977*

[30] *https://legiscan.com/MA/bill/S1841/2015*

[31] *http://billstatus.ls.state.ms.us/2015/pdf/history/SB/SB2676.xml*

		license to operate, for related purposes.	
Missouri	MO HB 924	[32]Allows testing of driverless motor vehicles until August 28, 2018	Pending –In House Committee on Transportation
New York	NY AB 31	[33]Provides for and regulates the operation and testing of motor vehicles with autonomous technology.	Pending –In Assembly Transportation Committee
North Carolina	NC SB 600	[34]Directs the Division of Motor Vehicles to study how to implement autonomous vehicle technology on the roads and highways of this state.	Pending –Passed Senate, In House Committee on Transportation
	NC HB 782	[35]Directs the Division of Motor Vehicles to study how to implement autonomous vehicle technology on the roads and highways of this state.	Pending – In House Committee on Rules, Calendar, and Operations of the House
North Dakota	ND HB 1065	[36]Provides for a legislative management study of automated motor vehicles.	Signed by Governor and filed with Secretary Of State
Oregon	OR SB 620	[37]Establishes process for certifying manufacturers for	Pending – In Senate Committee on

[32] *http://www.accessmissouri.org/bills/profile.php?id=28613*

[33] *https://www.nysenate.gov/legislation/bills/2015/a31*

[34] *http://www.ncleg.net/gascripts/BillLookUp/BillLookUp.pl?Session=2015&BillID=S0600*

[35] *http://www.ncleg.net/gascripts/BillLookUp/BillLookUp.pl?Session=2015&BillID=h782*

[36] *http://openstates.org/nd/bills/64/HB1065/*

[37] *https://olis.leg.state.or.us/liz/2015R1/Measures/Overview/SB620*

		purposes of testing, selling or operating autonomous vehicles on highways of state Prescribes vehicle and operator requirements for autonomous vehicles.	Business and Transportation
Tennessee	TN HB 616	[38]Prohibits local governments from banning the use of motor vehicles equipped with autonomous technology.	Substituted by SB 598
Texas	TX HB 933	[39]Creates a border security pilot program.	Pending –In House Transportation Committee
	TX SB 1167	[40]Relates to autonomous motor vehicles.	Pending – In Senate Transportation Committee
	TX HB 3690	[41]Relates to the operation of autonomous motor vehicles by the State Department of Transportation.	Pending – In House Committee on Transportation
	TX HB 4194	[42]Relates to the ownership and operation of autonomous motor vehicles	Pending –In House Transportation Committee

[38] .*https://trackbill.com/bill/tn-hb616-motor-vehicles-as-enacted-prohibits-any-political-subdivision-of-the-state-from-prohibiting-within-the-jurisdictional-boundaries-of-the-political-subdivision-the-use-of-a-motor-vehicle-equipped-with-autonomous-technology-if-the-motor-vehicle-otherwise-complies-with-all-safety-regulations-of-the-political-subdivision-amends-tca-title-4-title-5-title-6-title-7-and-title-55/1135614/*

[39] *http://www.legis.state.tx.us/tlodocs/84R/billtext/html/HB00933I.htm*

[40] *http://www.legis.state.tx.us/BillLookup/History.aspx?LegSess=84R&Bill=SB1167*

[41] *http://www.legis.state.tx.us/billlookup/text.aspx?LegSess=80R&Bill=HB3690*

[42] *http://www.aeltracker.org/bill-details/10368/texas-2015-hb-4194*

Utah	UT HB 373	[43]Modifies the Motor Vehicles Act by authorizing the Department of Transportation to conduct a connected vehicle technology testing program.	Signed by Governor
Washington	WA HB 2106	[44]Concerns autonomous vehicle testing in designated areas.	Pending – Carryover
2014			
California	CA AB 2258	[45]Authorizes the City of Lancaster to research and develop autonomous public buses.	Pending –In House Transportation Committee
Florida	FL SB 1272	[46]Requires the Department of State to consult with the Department of Law Enforcement in establishing a retention schedule for records generated by the use of an automated license plate recognition system. Authorizes the testing of such vehicles on certain roadways designated by the Department of Transportation and the applicable local government or authority.	Pending –In House Transportation Committee

[43] *http://le.utah.gov/~2015/bills/static/HB0373.html*

[44] *http://apps.leg.wa.gov/billinfo/summary.aspx?year=2011&bill=2106*

[45] *http://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201320140AB2258*

[46] *https://www.flsenate.gov/Session/Bill/2014/1272/Category*

Georgia	GA HR 1265	[47]Creating the House Study Committee on Autonomous Vehicle Technology	Failed– session adjourned.
	GA SB 369	[48]Defines "autonomous technology" and "autonomous vehicle." Requires autonomous vehicles meet federal standards, have means to engage and disengage the autonomous technology, have means inside the vehicle indicating when the vehicle is operating in autonomous mode, have means to inform the operator of a technology failure, and be capable of being operated in compliance with Georgia's uniform rules of the road.	Failed– session adjourned.
Hawaii	HI HB 1461	[49]Defines Autonomous Technology and Autonomous Vehicle. Authorizes the testing of autonomous vehicles on public roads, streets, and highways. Also establishes requirements for the operation of an autonomous vehicle.	Failed – failed first crossover deadline – second year of biennium.

[47] http://www.legis.ga.gov/legislation/en-US/Display/20132014/HR/1265

[48] *www.legis.ga.gov/Legislation/20132014/139361.pdf*

[49] https://legiscan.com/HI/research/HB1461/2015

	HI SB 1229	[50]Requires the Department of Transportation, in consultation with the insurance commissioner and the examiner of drivers of each county, to adopt rules providing for the operation of autonomous vehicles on state highways, including minimum safety requirements and any testing, equipment and performance standards that the department concludes are necessary to ensure the safe operation of autonomous motor vehicles on public roads. Also defines Autonomous Technology and Autonomous Motor Vehicle.	Failed – failed first crossover deadline – second year of biennium.
	HI HB 2420	[51]Requires the Department of Transportation to develop an autonomous vehicle testing program no later than January 1, 2015. Also requires the department to establish an application approval process by which manufacturers and other interested parties may test autonomous vehicles no later than January 1, 2016.	Failed – failed first crossover deadline – second year of biennium.

[50] *https://legiscan.com/HI/text/SB1229/id/916686*

[51] *https://legiscan.com/HI/text/HB2420/id/933292*

Louisiana	LA HB 937	[52]Authorizes the operation of autonomous motor vehicles.	Failed– session adjourned.
	LA HB 938	[53]Authorizes the research and testing of autonomous vehicles.	Failed– session adjourned.
Maryland	MD SB 773	[54]Prohibits an individual from operating an automated motor vehicle in automated mode on a highway. Establishes an exception to the prohibition by authorizing specified individuals to operate an automated motor vehicle in automated mode on a highway for research or testing purposes under specified circumstances. Requires an automated motor vehicle that is operated on a highway for research and testing purposes to have a specified individual present in the vehicle and to be properly titled and registered.	Failed
	MD HB 538	[55]Establishes the Task Force to Study Issues Related to the Use of Self-Driving Vehicles. Provides	Failed

[52] *www.legis.la.gov/legis/ViewDocument.aspx?d=880331*

[53] *https://lis.virginia.gov/cgi-bin/legp604.exe?161+sum+HB938*

[54] *https://mgaleg.maryland.gov/2014RS/fnotes/bil_0003/sb0773.pdf*

[55] *https://legiscan.com/MD/bill/HB538/2016*

		for the composition, chair and staffing of the Task Force. Requires the task force to make specified determinations, review specified information, consider specified issues, and make specified recommendations related to the use of self-driving vehicles. Requires the task force to report its findings and recommendations to the governor and the General Assembly on or before January 1, 2017.	
Massachusetts	MA HB 3369	[56]Defines Autonomous Technology and Autonomous Vehicle and authorizes the operation of autonomous vehicles without active control or monitoring by a human operator. Also requires autonomous vehicles meet federal standards, have means to engage and disengage the autonomous technology, have means inside the vehicle indicating when the vehicle is operating in autonomous mode, have means to inform the operator of a technology failure, and be capable of being operated in compliance with state laws	Failed– session adjourned.

[56] *https://malegislature.gov/Bills/186/House/H3369*

Minnesota	MN HB 1580	[57]Directs commissioner of transportation to develop legislation concerning autonomous vehicles.	Failed– session adjourned.
New Jersey	NJ SB 734	[58]Requires the New Jersey Motor Vehicle Commission to establish a driver's license endorsement for the operation of autonomous vehicles on state roads.	Pending-Assembly Transportation and Independent Authorities Committee
	NJ AB 1326	[59]Requires the New Jersey Motor Vehicle Commission to establish a driver's license endorsement for the operation of autonomous vehicles on state roads. Also requires the commission to establish regulations authorizing the operation of autonomous vehicles; including safety, testing, insurance, and registration requirements.	Pending-carryover
New York	NY SB 4912 NY AB 7391	[60]Defines Autonomous technology and Autonomous Vehicle and authorizes the testing and operation of autonomous vehicles on public roads. Also directs the Commissioner of Motor Vehicles to make	Failed– session adjourned.

[57] *https://legiscan.com/MN/drafts/HF1580/2015*

[58] *https://legiscan.com/NJ/text/S734/id/917652*

[59] *http://www.njleg.state.nj.us/2014/Bills/A1500/1326_I1.HTM*

[60] *https://www.nysenate.gov/legislation/bills/2013/S4912*

		recommendations upon additional legislative actions relating to autonomous vehicles.	
South Carolina	SC HB 4015	[61]Defines the term Autonomous Vehicle and other terms related to the manufacture and operation of an autonomous vehicle. Also provides a procedure for the testing and operation of autonomous vehicles in this state.	Failed– session adjourned.
	SC HB 4621	[62]Defines terms related to the manufacture and operation of autonomous vehicles and establishes requirements for the manufacture, operation, and testing of autonomous vehicles. Regulations include: autonomous vehicles must meet federal standards, have means to engage and disengage the autonomous technology, have means inside the vehicle indicating when the vehicle is operating in autonomous mode, have means to inform the operator of a technology failure, be capable of being operated in compliance with state laws, contain easily readable lettering reading	Failed– session adjourned.

[61] *https://legiscan.com/SC/bill/H4015/2013*

[62] *https://legiscan.com/SC/bill/H4621/2013*

		Self-Driving Vehicle and must have a special license plate as issued by the Department of Transportation.	
South Dakota	SD SB 139	[63]Authorizes the testing of autonomous cars on the highways of South Dakota.	Failed– session adjourned.
Washington	WA HB 1439	[64]Defines Autonomous Vehicle and allows their on-road operation. Also mandates the state patrol to adopt regulations governing autonomous vehicles that must include: a licensing requirement, compliance with federal standards, licensed driver responsibility for autonomous vehicle infractions, a restriction on commercial vehicles being autonomous, and a restriction that an autonomous vehicle must be a single vehicle only and be restricted from towing operations.	Failed– session adjourned.
	WA HB 1649	[65]Authorizes a vehicle controlled by autonomous technology to be operated on roads in the state by employees, contractors, or other persons designated by	Failed – session adjourned.

[63] *http://sdlegislature.gov/docs/legsession/2014/Bills/SB139P.htm*

[64] *http://apps.leg.wa.gov/billinfo/summary.aspx?bill=1439&year=2013*

[65] *http://apps.leg.wa.gov/billinfo/summary.aspx?year=2013&bill=1649*

		manufacturers of autonomous technology for the purpose of testing the technology. Also requires a person to possess a valid driver's license to operate a vehicle controlled by autonomous technology and submit proof of insurance or surety bond prior to the start of testing.	
Wisconsin	WI SB 80	[66]Authorizes autonomous vehicles to be operated on state highways if specific requirements are met. Also allows the DoT to promulgate rules establishing requirements for autonomous vehicles, including minimum safety and performance standards for these vehicles and for their operation.	Failed – session adjourned.
2013			
Arizona	AZ HB 2167	[67]Defines Autonomous Vehicles. Requires a person to have a valid driver's license to operate a vehicle with autonomous technology. Lists requirements for the autonomous vehicle, such as means to engage and disengage autonomous technology and an alert to	Failed – session adjourned.

[66] *https://docs.legis.wisconsin.gov/2013/proposals/sb80*

[67] *www.azleg.gov/legtext/51leg/1r/bills/hb2167p.pdf*

		notify the operator that some technology is failing. Allows autonomous vehicles to be operated on public roads for testing purposes only, provided there is proper insurance, surety bond or self-insurance acceptable to the Department of Transportation in the amount of $5 million. Directs the Department of Transportation to submit a report to the Senate President and Speaker of the House of Representatives by April. 1, 2015 recommending additional legislation or regulation.	
Colorado	CO SB 13-016	[68]Concerns the use of guidance systems to drive a motor vehicle.	Failed – session adjourned.
Hawaii	HI SB 1229	[69]Requires the Department of Transportation, in consultation with the Insurance Commissioner and the examiner of drivers of each county, to adopt rules in providing for the operation of autonomous motor vehicles on highways within the State.	Failed – failed first crossover deadline – second year of biennium.

[68] *http://www.leg.state.co.us/clics/clics2013a/csl.nsf/fsbillcont3/F6C2E6A3EE6EF24887257A920050A144?open&file=016_01.pdf*

[69] *https://trackbill.com/bill/hi-sb1229-autonomous-motor-vehicles-department-of-transportation-operation-and-licensing/568270/*

	HI HB 1461	[70]Authorizes for testing purposes the operation of autonomous vehicles on public roads, streets, and highways.	Failed – failed first crossover deadline – second year of biennium.
Michigan	MI S 169	[71]Defines Automated Technology, Automated Vehicle and Automated Mode. Expressly permits testing of automated vehicles by certain parties under certain conditions, defines operator, addresses liability of the original manufacturer of a vehicle on which a third party has installed an automated system, directs state Department of Transportation with Secretary of State to submit report by February 1, 2016.	Enacted and chaptered on Dec. 20, 2013.
	MI SB 663	[72]Limits liability of vehicle manufacturer or upfitter for damages in a product liability suit resulting from modifications made by a third party to an automated vehicle or automated vehicle technology under certain circumstances. Relates to	Enacted and chaptered on Dec. 26, 2013.

[70] *http://www.capitol.hawaii.gov/Archives/measure_indiv_Archives.aspx?billtype=HB&billnumber=1461&year=2013*

[71] *http://www.legislature.mi.gov/%28S%28hjqkn0cjlpdiy4eoomr24fhv%29%29/mileg.aspx?page=getobject&objectname=2013-SB-0169*

[72] *.http://www.legislature.mi.gov/%28S%28kstknf1vhezbvdmz5uvj51sg%29%29/mileg.aspx?page=getobject&objectname=2013-SB-0169*

		automated mode conversions.	
Minnesota	HF 1416 HF 1580	[73]Directs the commissioner of transportation to evaluate policies and develop a proposal for legislation governing regulation of autonomous vehicles by January 31, 2014.	Pending –In House Transportation Committee
Nevada	NV SB 313	[74]Requires an autonomous vehicle that is being tested on a highway to meet certain conditions relating to a human operator; requires proof of insurance; prohibits an autonomous vehicle from being registered in the state, or tested or operated on a highway within the state, unless it meets certain conditions; provides that the manufacturer of a vehicle that has been converted to be an autonomous vehicle by a third party is immune from liability for certain injuries.	Enacted and chaptered on June 2, 2013.
New Hampshire	NH HB 444	[75]Establishes a Committee of Legislators to study the use of autonomous vehicles in New Hampshire and instructs the committee to deliver a report by	Failed – Inexpedient to legislate

[73] *https://www.revisor.mn.gov/bills/bill.php?view=chrono&f=HF1416&y=2013&ssn=0&b=house*

[74] *https://legiscan.com/NV/bill/SB313/2013*

[75] *http://www.gencourt.state.nh.us/legislation/2013/HB0444.html*

		November 1, 2013.	
New Jersey	NJ SB 2898	[76]Directs Motor Vehicle Commission to establish driver's license endorsement for autonomous vehicles.	Failed – session adjourned.
Oregon	OR HB 2428	[77]Defines Autonomous Vehicles. Allows a manufacturer to test an autonomous vehicle on public highways provided the manufacturer has a certificate of approval from the Department of Transportation. Directs the Department of Transportation to establish standards for equipment and performance of autonomous vehicles in order to ensure the safety of autonomous vehicles on state highways. A manufacturer must provide proof of liability insurance not less than $5 million. Lists various requirements that the autonomous vehicle must include, such as interior indicator lights, a failure notification signal, etc. Requires the operator of an autonomous vehicle to hold a certain driver's license class for the autonomous vehicle.	Failed – session adjourned.

[76] *https://legiscan.com/NJ/bill/S2898/2012*

[77] *http://gov.oregonlive.com/bill/2013/HB2428/*

Texas	TX HB 2932	[78]Relates to the operation of autonomous motor vehicles. Defines Autonomous Vehicle and Autonomous Technology. Authorizes a person who possesses a valid driver license to operate an autonomous vehicle. Directs Department of Transportation to establish standards for equipment and performance.	Failed – session adjourned.
Washington D.C.	DC B 19-0931	[79]Defines Autonomous Vehicle as a vehicle capable of navigating District roadways and interpreting traffic-control devices without a driver actively operating any of the vehicle's control systems. Requires a human driver prepared to take control of the autonomous vehicle at any moment, restricts conversion to recent vehicles and addresses liability of the original manufacturer of a converted vehicle.	Enacted and effective from April 23, 2013.
2012			
Arizona	AZ HB 2679	[80]Directs the Director of the Department of Motor	Failed – session adjourned.

[78] *https://legiscan.com/TX/bill/HB2932/2013*

[79] *http://openstates.org/dc/bills/19/B19-0931/*

[80] *http://www.azleg.gov/legtext/50leg/2r/bills/hb2679p.htm*

		Vehicles to adopt rules authorizing the operation of autonomous vehicles on highways in the state, including minimum safety and insurance requirements. Restricts testing to certain areas and provides for autonomous vehicle drivers license endorsement.	
California	CA SB 1298	[81]Requires the Department of the California Highway Patrol to adopt safety standards and performance requirements to ensure the safe operation and testing of autonomous vehicles, as defined, on the public roads in this state. Permits autonomous vehicles to be operated or tested on the public roads in this state pending the adoption of safety standards and performance requirements that would be adopted under this bill.	Enacted and chaptered on September 25, 2012.
Florida	FL HB 1207	[82]Authorizes the operation of vehicles equipped with autonomous technology for the purpose of testing the technology. Also directs the Department of Highway Safety and Motor Vehicles to prepare a report relating to the safe operation of	Enacted and chaptered on April 16, 2012.

[81] *https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201120120SB1298*

[82] *https://www.flsenate.gov/Session/Bill/2012/1207*

		vehicles equipped with autonomous technology on public roads.	
	FL HB 599	[83]The relevant portions of this bill are identical to the substitute version of HB 1207.	Enacted and chaptered on April 29, 2012.
	FL SB 1768	[84]Directs the Department of Highway Safety and Motor Vehicles to prepare a report relating to the safe operation of vehicles equipped with autonomous technology on public roads, to be submitted no later than February 1, 2014.	Substituted by HB 1207 on March 9, 2012.
Hawaii	HI HB 2238	[85]Authorizes the issuance of a driver's license for operation of autonomous motor vehicles; requires the Department of Transportation to adopt rules for the operation of autonomous motor vehicles.	No longer relevant to autonomous vehicles
	HI HCR 212	[86]Urges the Department of Transportation to review and report on policies relating to the use of driverless vehicles in Hawaii.	Failed – session adjourned.

[83] *https://www.flsenate.gov/Committees/billsummaries/2012/html/210*

[82] *https://www.flsenate.gov/Session/Bill/2012/1768*

[85] *http://www.capitol.hawaii.gov/session2012/bills/HB2238_.HTM*

[86] *https://legiscan.com/HI/text/HCR212/id/619422*

	HI HR 163	[87]Identical to HI HCR 212	Failed – session adjourned.
New Jersey	NJ AB 2757 NJ AB 3020	[88]Defines Autonomous Vehicle. Directs the New Jersey Motor Vehicle Commission to adopt rules for driver's license endorsement and requirements for operation of an autonomous vehicle, including insurance, safety standards and testing.	Failed – session adjourned
Oklahoma	OK HB 3007	[89]Directs the Department of Public Safety to establish an endorsement for the operation of autonomous vehicles on highways. Also directs department to adopt rules relating to autonomous vehicles, setting forth requirements for operating such a vehicle in the same, including safety and insurance requirements	Failed – session adjourned.

Tab.3.2 Introduced Autonomous legislation's timeline

Source: Anne Teigen and Amanda Essex

Every State has enacted its own laws and everyone decides when they are out-of-time and they need to be checked. That complicates a common framework and even some of them ban the use of Autonomous. These differences are shown as follows in Fig.3.2.

---

[87] *http://openstates.org/hi/bills/2011%20Regular%20Session/HR163/*

[88] *https://legiscan.com/NJ/bill/A2757/2012*

[89] *https://legiscan.com/OK/bill/HB3007/2012*

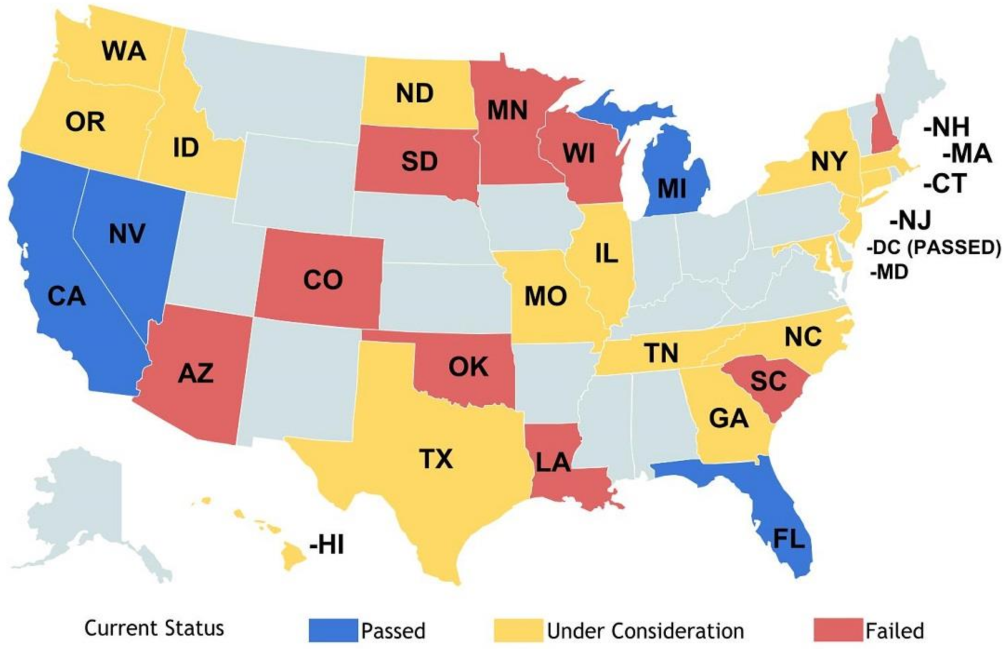

Fig.3.2 States which have introduced autonomous legistaliton

Source: Gabriel Weiner and Bryant Walker Smith

Even though every year more States are allowing both testing and operation, we are still far away from a common framework. Legislation evolves going hand in hand with Autonomous development and that can explain the different velocity as regards enacting laws. Those States with companies demanding news legislations are going to be the reference in the next years. As I said before, Autonomous companies want to be supported by reliable legislations to keep improving their technologies. We should not forget that, having a competent legal framework, have an effect on attracting companies and then on their economies. The own legislation of each State is also a problem to face up to. In a near future, the United States should unify their laws to allow the Autonomous in all the country and then they could become the reference as to legislation.

3.2 European Union

Now in Europe, we also find Autnomous legislation, but as in the United States, there are several differences between the members. The regulations on how new vehicles should operate and be designed are harmonized within the EU through Framework [90]Directive 2007/46/EC. The purpose of the regulations is to create an internal market within the

[90] *http://ec.europa.eu/growth/sectors/automotive/legislation/motor-vehicles-trailers/directive-2007-46-ec/index_en.htm*

Community and aims to ensure a high level of road safety, health protection, environmental protection, energy efficiency and protection against unauthorized use. In other words, the EU regulates which requirements have to be met. However, the more detailed technical provisions are mainly prepared within UNECE [91]WP 29 and can be found in the UNECE regulations to which the EU legislation refers. There is a certain degree of scope for guaranteeing national, alternative requirements and permitting exceptions for test operations, for instance.

The public authorities in the Union have also been active in implementing and further developing innovative concepts for automated driving. Even having a common legal framework referring to road legislation –and therefore also for autonomous vehicles– every single state has both the responsibility and duty to enact its own laws. That is why, even the above referred common framework, it varies depending on the considered country. However, there are approaches that represent magnificent progress in this field. Now it is shown an overview of the enacted legislation in the different countries and the different initiatives promoted to enable the use of driverless cars.

3.2.1 Netherlands

The first step of the Dutch government was announced in June 2014, when it issued its intention to allow large-scale testing of self-driving vehicles on Dutch roads, and the necessity to amend the actual legislation to such activities. The private sector is also concerned with this hot topic. This sector has been carrying out small-scale trials on the Dutch roads, as a part of the Dutch Integrated Test Site on Cooperative Mobility (DITCM) in Helmond.

The next step was in January 2015, when a proposal to extend exemption rules to allow the large- scale testing of self-driving cars and trucks was approved. The testing started in summer 2015 under new legislation approved by the Dutch Parliament to allow these trials.

On behalf of the minister of the Ministry of Infrastructure, The Dutch Vehicle Authority (RDW) has published an exemption process for large-scale testing on public roads. It consists of the following steps, being the most relevant the last one:

1. *Intake: "Mission analysis" functional description, comprehensive risk analysis and EMC.*

2. *Desk Research, in consultation with the applicant.*

3. *Testing on a closed proving ground: physical inspection of the vehicle(s), "Happy flow" testing and stress testing.*

[91] *http://www.unece.org/trans/main/wp29/introduction.html*

4. Limited admittance with an exemption: requirements and conditions such as insurance, test times, test location, duration, monitoring and logging.

Now in 2016, the Dutch will occupy the presidency of the European Union and have stated their intention to continue developing Autonomous legislation and solve future problems. They also want to take advantage of their privileged position in the EU government to initiate amendments to international regulations and have already launched a study into potential issues such as liability, driving skills requirements, traffic data and the possible impact of such vehicles on infrastructures.

3.2.2 United Kingdom

As part of the 2013 National Infrastructure Plan, the government pledged a [92]review of the legislative and regulatory framework to enable the trialing of driverless cars on UK roads. On 30 July 2014, the government also launched a driverless cars [93]competition inviting different UK cities to join together with businesses and research organizations and host vehicle trials locally. The results were announced in December 2014 with Greenwich, Milton Keynes, Coventry and Bristol being selected, and £19 million being provided by the Government to allow testing of automated vehicle technology.

The UK Government also published a detailed review of [94]Regulations on 11 February 2015, examining the regulatory framework for the safe testing of driverless cars. The main conclusions were:

- Driverless vehicles can legally be tested on public roads in the UK today, providing that a test driver is present and takes responsibility for the safe operation of the vehicle, and ensuring that the vehicle can be used compatibly with the road law.

- A Code of Practice was published during the spring of 2015 for those wishing to test driverless vehicles on UK roads.

- Need to review and amend domestic regulations by the summer of 2017 to accommodate driverless vehicle technology.

- Liaise at an international level with an aim to amend international regulations by the end of 2018.

[92] *https://www.gov.uk/government/consultations/driverless-cars-regulatory-testing-framework*

[93] *https://www.gov.uk/government/news/uk-government-fast-tracks-driverless-cars*

[94] *https://www.gov.uk/government/publications/driverless-cars-in-the-uk-a-regulatory-review*

3.2.3 Sweden

The Swedish Road Traffic Ordinance still contained limitations about the use of driverless cars because of the lack of a driver and permitted limited exception to general provisions. Due to the limitations in autonomous legislation, specific driver license requirements and liability rules, the Swedish Transport Agency published in May 2014 a report, in which they insist on the need of revision of such areas, to allow testing of vehicles with high levels of automation.

Current legislation allows the use of vehicles with high degree of automation, similar to driverless cars, with testing purposes into Swedish roads. Due to this, at the moment the road traffic legislation is not an impediment to such technology. In addition, the Swedish Transport Agency is able to grant exceptions, to those vehicles which do not fulfill with the obligatory technical requirements due to their innovative technology, to allow their circulation. At the moment, the Swedish industry is developing a driverless car test, called Drive Me, through its roads.

3.2.4 Italy

The use of autonomous vehicles in Italian roads is totally banned for safety reasons. However they exceptionally allow some test on short road section, once it has been secured to prevent traffic accidents. Regarding to autonomous technologies and in particular with automated transport systems, they may be considered legal if they are certified according to a technical standard developed for rail systems.

There are only two papers which provide information or need to amendment of the current autonomous legislation. The first one is the Decree of 1 February 2013 on the diffusion of Intelligent Transport Systems (ITS). It was published in the Official Journal of the Italian Republic and signed by the Minister of Transport. The second one was approved on 12 February 2014 by the Minister Maurizio Lupi, regarding to the National Action Plan for Intelligent Transport. The problem is that both mentioned documents deal exclusively with cooperative driving, whereas no mention was made of automated driving.

3.2.5 France

Due to its limitation for being a signatory of the Vienna Convention, France has been pressing for its amendment and, in July 2014, the France Government published its own roadmap. On it they stated which roads allowed the use of automated vehicles. They also included proposals such as pilot zones for testing, changes to driver training, Research and Development projects running to 2018 and the development of regulatory requirements to support testing automated vehicles and their entry to the market.

3.2.6 Germany

The Federal Highway Research Institute published during the 2012 a report in which it summarized the situation between Autonomous and the, at that time, current German legislation. The conclusions of that report were that the legislation was compatible with the existing automation levels, but that consideration does not included neither highly automation (level 4) nor full automation (level 5). The German legislation only guaranteed the right regulation to that levels in which the driver has constantly the control over the vehicle (from level 0 to 3).

Later, a meeting about automated driving took place in November 2013. The participating of that round table were the main stakeholders involved in that task. There were different representatives, such as Federal Ministries, public authorities, autonomous industry representatives, insurance companies, user associations, technical inspection and research institutes. Some work groups were created according to different tasks as Legal Issues, relation Driver/Vehicle and Research. The main objective of the meeting was to define a common encounter point between all the stakeholders and driverless cars, creating so the basis of the engagement within Autonomous and society.

The German legislation states that research and development activities cannot be prohibited, so it allows the use of Autonomous for such purposes. each federal state in Germany can grant exemptions from the technical requirements of the German Road Traffic Licensing Regulations. That In addition, means that every autonomous vehicle, under the supervision of an authorized responsible, is legally allowed to operate trough German roads. Thanks to that, some tests have been performed during the last years. Moreover, the German transport Minister announced in January 2015 that an autonomous test will take place across the A9 autobahn between the cities of Munich and Nuremberg.

3.2.7 Spain

The Spanish legislation is still restricted by both the Vienna and Geneva Convention. This is still visible at the Spanish Road Code, which contains the statement: "*Drivers should at all times be in a condition to control their vehicles*". Even that statement could be problematic for autonomous vehicles it does not ban explicitly the use of them, but require the supervision of a human during its operation.

The public authorities are also active regarding to development of autonomous technologies. The Spanish Ministry of Science and Innovation funded diverse projects focused on the development, implementation and validation of methodologies for supporting a system control and navigation of automated vehicles. Some platooning trials have been undertaken, and they plan to undertake new test in autumn 2016. Spain also hosted the SARTRE FP7 project for open-road platooning. FP7 projects Citymobil and Citymobil 2 hosted demonstrations of automated transport in Castellon, León and San Sebastian. The Spanish government has also supported and allowed the traffic of an automated vehicle on 100 kilometers of a highway without requiring input from the driver.

3.2.8 Belgium

The Government performed a few trials during the last year on the Flemish roads. In addition, the governmental department of Mobility has also announced a partnership agreement with the autonomous company Tesla Motors, to allow their testing into its highway network.

3.2.9 Finland

During the 2015, the Finnish government was preparing experimental legislation to allow the use of automated vehicles within the next five years. They have guaranteed their legal use into restricted areas and at certain times, only for testing purposes and under special traffic permit.

3.2.10 Greece

During the 2013, and taking advantage of the celebration of CITYMOBIL2 pilot into Greek lands, the Ministry of Infrastructure, Transport and Networks, in cooperation with stakeholders in the country, began discussions about the allowance of autonomous vehicles into their roads. In May 2014, a meeting between the different European countries took place in Athens with the aim of tracing the basis of driverless cars legislation. The participating countries were Greece, France, Italy, Finland, Spain, Sweden, Germany, UK, Poland, Malta and Cyprus. As a result, a legislative proposal was written and is still pending to the Parliament acceptance. It group together different initiatives such as proposed routes, technically qualified staff monitoring and under local authorities supervision.

3.2.11 Switzerland

The Federal Department of Environment, Transport, Energy and Communications in Switzerland (UVEK), allowed during the spring of 2015 the first trial of an autonomous vehicle into their roads. They allowed the free circulation of a driverless Volkswagen Passat under the supervision of an on board driver. The test was performed by Swisscom through the streets of Zurich.

3.3 Australia

With the goal of improving freight access and productivity, the State Road Authority in Victoria (also known as VicRoads) has released an industry framework for trialing road freight intelligent systems on Victorian roads. The stakeholders of such proposal were the transportation industry, in particular the logistic sector. They highlighted the relevance of this sector through ITS by improving efficiency, reliability and safety. Both the Victorian Government and the State Road Authority are concerned to facilitate and stimulate the innovation in this industry, and due to this they are working together to develop new legislative standards to support the research and development.

On his behalf, the South Australian Government has stated its intentions to allow autonomous cars into their roads, urging the different involved parties to check and develop new legal frameworks. They also have planned to introduce new legislation to State Parliament during the present year.

3.4 Japan

In May 2014, the Japanese Government launched a Cross-Ministerial Strategic Innovation Promotion Program (SIP), which contains a research section about Automated Driving systems. Such technologies were recognized as a way to reduce traffic congestions and accidents, helping to improve the everyday life of cities. It also recognized development and proof of automated driving systems. The Japanese Minister has also encouraged European authorities to further the development of such technologies. The Vice-Minister of Land, Infrastructure, Transport and Tourism has also been active promoting Autonomous initiatives. Due to this, this ministry presented a Study Panel for Autopilot Systems to discuss with car manufacturers which are the main challenges to face up to develop autonomous technologies. Japanese car manufacturers are in steady progress too, specially Toyota which is considered one of the referents relative to Autonomous.

Today, Japan, and more specifically Tokyo, is considered to be the first city to allow the free use of autonomous vehicles into their urban roads. The government is decided to introduce the new generation of urban transportation systems with the objective to be totally implemented during the 2020 Olympic Games.

3.5 China

Due to China is not a signatory country neither the Vienna nor the Geneva Convention, is not possible to find any law against the use of Autonomous into Chinese road. Although it may not appear so, any person or organization wishing to test driverless car is only obliged to be under Chinese driving license fulfillment and provide a Chinese number plate to the car. Even this favorable scene, and because of lack of privacy around autonomous patents, this country is not considered as a model where to develop such technologies.

3.6 Singapore

The Land Transport Authority of Singapore (LTA) signed with the Agency for Science, Technology and Research (A*STAR) a five-year Memorandum of Understanding. This partnership was called the Singapore Autonomous Vehicle Initiative (SAVI), and its main objective was the creation of a common association with the duty to develop and implement autonomous vehicles, mobility and road systems. Another objective was to perform some trials, trying to involve both public and manufacturer authorities. The LTA took the role of regulator with the aim to allow progressively autonomous cars into their transport network, while A*STAR was held responsible of create an Autonomous roadmap and keep developing new technologies.

The first trial took place in January 2015 in a 200-hectare enclosed area. The only requirements required for such purposes were the availability of safety procedures, to take control over the car at any time, for instance, and third-party insurance.

3.7 South Korea

Research facilities in Korea distinguish autonomous cars between, on the one hand which obtain information through their own sensors and, on the other hand, which combine the information coming from their sensor and from the own infrastructure –considered as an Automatic Vehicle Guidance System (AVGS)–. Recently the Electronics and Telecommunication Research Institute (ETRI) has started research on autonomous vehicles by further developing its IT Convergence technology towards AVGS.

With the aim to decrease the number of deaths in traffic accidents and to improve the circulation into their roads, the Korean Ministry of Land, Infrastructure and Transport (MOLIT) created a round table to develop safety standards in the transport sector and has promoted the Research&Development of autonomous transportation technologies until 2016.

The information above presented is a summary of the paper written by HOTTENTOT C. ET AL. [2015], in which they show those countries that are concerned about the use and development of autonomous vehicles. However, this paper is focused only in legislation, as opposed to the latter that speaks about autonomous vehicles in general and without going far into autonomous vehicles legislation. Regarding the legislation above mentioned, almost all the autonomous-leading companies are being weighted down by the restrictive legislation present in many countries, when they try to test such innovative technologies on real motorways. As noted in previous chapters, this is because many of them are signatories of the Geneva or Vienna Convention, which enact that *every moving vehicle or combination of vehicles shall have a driver*. This fact represents a legal barrier to Autonomous' developers when they try to test such technologies under real traffic scenarios. Nonetheless, every country has specific articles in its own legislation to ensure and defend the exercise of the right of inquiry, and consequently the authorization to test driverless cars into their road network for trialing purposes, provided that they fulfill with specific and restrictive driving rules. All these circumstances contribute to complicate, even more, the legal fitting of Autonomous into society, against which all the stakeholders involved have to deal with. Moreover, these parties are not only the own autonomous manufacturers, but also the Government Administration and other involved agencies as insurance companies or other automobile associations. Due to the fact that such technology is conceived to a better social development, avoiding accidents and pollution, all the parties have to be involved and work together and hand in hand in favor of this technology. In addition, this blended approach would ease the resolution to future problems, including the legal framework, in a better and faster way.

As noted, the legal problem coming from the fitting of Autonomous in the Society is not the only one, being likely to find additional problems. This is the reason why, throughout the next section, we will see the main problems with which Autonomous will have to face, in order to become a real technology.

4 Issues

One might think that the only challenge of autonomous vehicles in a mid- to long-term is going to be legislation and how it regulates the use of such vehicles. Additionally, it is true that when Autonomous become a reality, and due to sever road regulations, the fitting between both parties is going to be a major problem. However, it is not going to be the unique problem that such innovative technology will have to deal with. Today, it is likely to find additional challenges, which have to be monitored and analyzed in detail with the goal of guaranteeing the future success of Autonomous. Then, the purpose of this section is to identify the main challenges that surround autonomous vehicles and could complicate their development.

4.1 Technological Challenge

Today, changes are occurring on motor vehicle transport, due to the transition from vehicles driven by humans to self-driving vehicles, known also as Autonomous or driverless cars. The technology used in the latter allows the driver to do different activities during the travel, never seen before. It will be possible thanks to the autonomous intelligence or predefined driving algorithms, from which the car will know when it has to accelerate, brake or even overtake other vehicles.

But such innovative technology is not a concept of near future, because it has been used during the last decades. However, the important thing here is the known as level of automation, which could generate technological problems for those highest (semi-autonomous or autonomous level). For example, air bags or anti-blocking systems are considered one of the first semi-autonomously functioning systems, and today they are vastly used. More current examples are the parking assistance systems or lane departure assistants.

Even the large track-record of semi-autonomous technology, the really challenge arises for the fully autonomous level. Such level of automation allows the driver, as seen before, to do different activities during the travel, without keeping an eye on driving. Consequently the Autonomous has to be able to face all the current and unusual situations happening during the travel as best as it can. For this reason, and knowing that human lives may depend on the decision made by the Autonomous, this technology has to be tested and validated with regard to high quality and safety requirements.

4.2 Constitutional Provisions

Due to the characteristic rigidity of Constitutions, they do not consider terms referred to technical innovations and do not establish how the system has to deal with new technologies. For example, constitutional texts do not recognize the use of terms as “technical equipment”, "technology" or "technical progress", and do not include any provision on managing technology.

Nevertheless, the fact that Constitutions do not considerer autonomous terms does not imply that they are not allow. In addition, almost all Constitutions have specifics articles to cover and guarantee research and development activities. For example, the Article 5 of the German Constitution contains the guarantee of freedom to conduct research, and the Article 14 the protection of new technologies as a kind of a paten right.

According to HILGENDORF E. [2015], it is easy to find several constitutional Principles that can be applied to ensure and protect the different research and development activities in the field of autonomous technologies. They are as follow:

> *- The principle of research is free and constitutionally protected.*
>
> *- In case of injury, rules provide for adequate compensation to be granted.*
>
> *- The description of motor vehicle transport as “safe and efficient” expresses particularly clearly the two necessary conditions which the state must ensure.*
>
> *- Because of the inherent hazard potential of vehicles, research on autonomous vehicles has always been subject to legal constraints, aimed at reducing possible risks created by the new vehicles.*

The protective duty of the state with respect to potentially harmful effects of new technologies may be derived mainly from the objective guarantees of fundamental rights. This concerns directly the technologies with a high probability of causing injury, such as Autonomous.

4.3 Road Traffic Law

Due to several countries are signatories of the Geneva or Vienna Convention, their road traffic legislations are also based on their articles. In case of Europe its road traffic legislation is largely based on the Articles stated by the Vienna Convention, two of whom are specially restrictive with the use of autonomous driving systems. These two Articles are as follow:

> - Article 8 stipulates *that every moving vehicle or convination of vehicles must have a driver.*
>
> - Article 13 states that *a driver must be in control of his vehicle in all circumstances*.

The implementation of these two provisions means that every vehicle should be at all times under control of a human driver, who must be prepared to take control of the vehicle.

But during the spring of 2014 several European countries insisted on the necessity to amend the stipulations of the Vienna Convention. This was seen as a breakthrough towards the regulation and approval of autonomous vehicles. The gist of the scheme is that driver assisted systems, which are already accepted by a specific UN body, would also be deemed to meet the requirements of Articles 8 and 13. It should be pointed out that the Vienna Convention, conceived as an international agreement, cannot be applied to motorist and car manufacturers in any European country.

Despite the European regulatory framework, every country is responsible of the admissibility of autonomous systems into its roads. This is because the allowance of such vehicles is regulated by the provisions of national law contained in specific national road traffic acts and highway codes. That also means that an amendment to the Vienna Convention does not imply any change in national road regulations. Only until the amendment become effective nationally, the agreement between autonomous vehicle and legal specifications could not be achieved on the basis of a broad and technology-friendly interpretation of the current regulations.

4.4 Criminal Law

Regarding criminal liability, the driver is going to be the primarily responsible in case of incurring in criminal charges, but both the seller and the manufacturer would have to face criminal liability. Here could be of application the offense of causing bodily harm by negligence (e.g., under § 229 German Criminal Code, StGB).

A common example about what has been said before could be a person hit by a car, while the driver was using the parking assistant system. All the elements of the offense would be fulfilled, namely an “action” (allowing the computer to park the car), “result” (the injury to the person) and “causation” (if not for the parking, the person would not have been injured). But the question appears when asking whether the driver acted negligently parking the car. Negligence applies to a person when, he or she is not enough careful. So the problem is to define what really constitutes reasonable care for the driver. Therefore, in the example discussed above, the driver should have checked whether there was anything at the parking space and supervise the manoeuvre at all time.

But, whether the autonomous technology is conceived to drive laid-back, why should the driver look after driving? The answer is that the driver is the main responsible for the proper function of the autonomous system, and whether necessary, taking the proper corrective actions. As a consequence, the driver should be focused on driving instead of reading the newspaper or writing an email, reducing substantially the advantage of such autonomous technologies. This is also known as the “dilemma of control”.

That is why in case of fatality, the criminal liability between the different parties involved could be not well defined and that supposes a problem. Moreover, both the seller and the manufacturer could have to face legal responsibilities, being vicariously responsible, in case of obvious and proven failure.

4.5 Safety and Liability

In case of traffic accident, is really important –for both road users and insurance companies– to determine which party is responsible of the happened occurrence, either the driver of the vehicle, the owner or the manufacturer of the Autonomous. Until these three parties define their extent of liability in a well-adjusted way, autonomous vehicles and their use will be not accepted by society. An alternative solution proposed by HILGENDORF E. [2015] is to isolate the victim –in case of accident– from the matter of guilt, what means indemnifying the first without regard to who is the responsible. Then the matter of guilt becomes secondary, what implies that an exhaustively analysis in insurance's field has to be performed to solve the problem of determination of responsibilities.

Finding the best way to solve the problem mentioned above will provide impetus to the acceptance of autonomous vehicles, considering that such innovative technology has to be introduced gradually in a mixed scenario together with non-automated vehicles. Whether we consider a future scenario with multimodal ways of transportation –including Autonomous–, the liability of automated driving has to safeguard both public services, roads users and pedestrians in order to ensure the fair safety of all road users.

With the purpose of maintaining security in roads, autonomous vehicles will have to deal with three main challenges. The first one is their communication with those vehicles that do not use autonomous technology. The second one is how these vehicles are going to be able to identify vulnerable road users (e.g., cyclists or pedestrians), considering that in sudden and unexpected cases their identification may be not effective and then, they could be not detected by autonomous vehicles. The third and major challenge is the regional dependence between multimodal transport and Autonomous. For example, the use of bicycles depends on the country, and even the city, so autonomous driving will have to face with several and changeable requirements. The solution to the latter is not easy to determine and will slow down –even more– the implementation of Autonomous in real traffic.

4.6 Civil Liability

As reported by HILGENDORF E. [2015], liability for possible damage occurring in the road transport context is primarily liability in tort (e.g., under §§ 823 German Civil Code, BGB), which under certain circumstances may be augmented by contractual liability. Additionally also exists special liability to cover road traffic, under which the owners of vehicles may incur liability, which for drivers is supplemented by a specially devised form of fault based liability.

The above mentioned rules are specially relevant due to they apply regardless of whether the vehicle that has caused damage is Autonomous or not.

In reference to this, HILGENDORF E. [2015] states that there is no need for change required. The use of autonomous technology would contribute to decrease the percentage of road accidents and therefore the number of fatalities. However and as a consequence of the subsidiary liability of manufacturers, the number of cases in which they could face with a lawsuit, rather than the owner or the driver, would increase largely. This is because the functioning of the autonomous system could be seen as the cause of the accident, rather than the behavior of the driver, and as a consequence the liability for failures of the autonomous system would be transferred to the manufacturer. In addition this could be seen as a problem by insurance companies.

4.7 Liability of Providers

Despite the obvious liability of providers in case of accident, there has been no much discussion about their role in the future ways of transportation. In addition, and due to the increase in vehicles' automation levels, the data exchanged between such vehicles and their surrounding is also going to increase, and consequently their exposure to external risks. For example, it would be usual that autonomous systems download information, either for traffic or weather information purposes, from servers that could be affected by malicious software. Then, through this software third parties could take control of the *non-so-autonomous* vehicle, blocking the vehicle, changing it setting or even causing an accident. So that is why liability of providers is specially relevant, due to they could have to face criminal liability in case of fatality.

For those cases, in which Internet is used as a tool to commit a crime, some questions emerge in relation to which party is really ultimate responsible, whoever is giving Internet access to the criminals, or whoever has the malicious software on its servers. In connection with the example mentioned above, the person who set up the software would be the responsible party in case of damage. HILGENDORF E. [2015] states that a partial solution could be deduced from the European E-Commerce Directive, in which *most national telemedia acts (and similar provisions) stipulate that whoever stores their own data on their own servers and makes it available for users, incurs unlimited liability*. In addition, *whoever stores third party data on their own server and make it available for users, only incurs liability whether the individual most certainly knew that the stored data was infected by malicious software.* Just providing Internet access to unlawful contents does not categorically implies incurring liability, except in case of favoring the distribution of such content.

The criterion mentioned above is applicable for both autonomous vehicles as wells as to that information obtained from outer servers. In addition, is still necessary to determine whether the privileged status characteristic of providers, will protect them in case of a criminal behavior carried out by one of their customers.

4.8 Data Protection Law

Due to the amount of data that autonomous vehicles will have to exchange with other vehicles or servers, and knowing that such data could contain personal information (e.g., routes, schedules or even the behavior of the “driver”), data protection is going to be another challenging issue to cope with. In addition, such increase in Autonomous will also grow the necessity for reliable data, putting additional pressure on data protection and from which sort of sources is it obtained.

As seen before, this data will probably contain personal data due to the data required by the vehicle is, ultimately, a request of the driver and as a consequence it contains behavioral patterns and tastes. The problem is that today it is unclear whether Autonomous are going to store all this driver-generated data or whether quite the opposite, are going to be third parties (e.g., manufacturer, supplier or seller) and the obvious defenselessness of the driver. In addition, another question arises when talking about data protection and it is who are going to be able to access to the stored data, and their use of such information.

The access to such data can be advantageous –and even profitable– to third parties. These parties could be both private (e.g., insurance companies, manufacturers and for companies specialized in drivers’ preferences) and public (e.g., Department of Transportation or councils). It is also necessary to underline, that almost all data generated by humans contains personal information, and this is going to be an upward trend in the mid- and long-term. According to HILGENDORF E. [2015] the practice of “big data” in such context contradicts the principles of the European Data Protection Law (e.g., against *minimization and purpose limitation of data*). Therefore the acceptance of autonomous vehicles will deeply depend on fulfillment of the European Data Protection Law.

4.9 Cyber Security

Another important challenge for autonomous vehicles is cyber security, which is owing to their high level of connection, typical of such innovative technology. As is known, Autonomous will be able to communicate with other vehicles or external servers, most of the time through Internet. But that is only the positive view, considering that on the other hand Internet is also an open source of risks. One example are hackers, who could take control over the Autonomous’ steering, thus creating chaos.

In line with this, [95]Hugh Boyes –the cyber security lead at the Institution of Engineering and Technology in London– stated that *the reliability and security of software used in driverless cars will be a major cause of concern for both manufacturers and insurers, and if hackers found a way to target these vehicles, this could present a whole new set of challenges that the industry is not currently equipped to tackle*. In addition, and due to the current and updated new ways of terrorism, Autonomous could be in the spotlight of such criminals as a tool to commit future crimes.

Due to this risk and as a counter-measure, autonomous manufacturers have to design specific firewall to increase their security against cyber attacks, ensuring their "fail safe". Some examples could be manual braking systems to avoid an electronic failure, or the "disconnection" of the autonomous vehicle from the network, reducing to a minimum the possibility of cyber attacks. In case of testing, manufactures should have to ensure that all prototypes tested have high levels of security, including that they could function manually –by remote control or by computer– in case of failure of the autonomous system.

4.10 Ethics and Algorithms of Death

The last important issues regarding autonomous vehicles are ethics and algorithms of death. Due to the technology used, most of the time the own vehicles will have to make its own decisions, some of which are going to have –unavoidably– ethic features. This means that, even Autonomous are conceived and designed to be safer than non-autonomous vehicles (e.g., reducing the number of accidents and traffic fatalities) more than likely they are going to be involved in accidents. And then, in case of accident, every Autonomous will have to be able to select the path with the lowest damage or likelihood of collision. So, Autonomous would have to "decide" who will die –in case of fatal accident–, being that the ethic challenge. For this reason, manufactures will have to deal with such complex scenarios with the final purpose of reducing the number of victims.

As seen before, the main problem appears when the autonomous system has to carry out, by itself, maneuvers in special conditions (e.g., a sudden accident). In case of accident, autonomous systems are able to decide, depending on the inputs received by its sensors and the modeled scenario, which path has the lowest damage or likelihood of collision in a faster and efficient way than human drivers. As an example, imagine that an autonomous vehicle is travelling on the highway, and unexpectedly a motorcyclist goes across its lane. Then the dilemma arises: the vehicle can maintain its trajectory injuring the motorcyclist or swerve to the right damaging both vehicle and guardrail. In this case the principle of causing the least harm applies, being the latter the best and chosen option due to causing damage to

[95] *https://www.uschamberfoundation.org/blog/post/were-ready-driverless-cars-what-about-new-cyber-threat/43243*

property has a lower value than to human lives. For this reason autonomous vehicles will have to be programmed to choose the least harmful scenario.

In the previous example the decision is fairly clear, but the problem comes up when both choices involve personal injury and one must be weighed against the other (e.g., deciding whether injuring a child or an old man, or three persons against five). Some legislation state that *human live cannot be weighed against one another and that every human life represents the highest value so that one life cannot be worth more than any other.* This article is a statutory right included in the German and Swiss Legislation. The problem is that human can justify actions on a case-by-case basis according to instant moral judgment but autonomous vehicles or computer machines can only decide according to their preset programming or algorithms of death.

5 Proposal

Once identified and characterized the main problems and challenges that autonomous technology will have to face, along with the absence –in almost all countries– of Autonomous legislation, the main objective of this section is to lay the foundations for a continuous development of such technology. Due to this technology is really innovative and never seen before, it has to face several complex problems. As a consequence, the first and main objective of this section is the definition of a legal framework that takes into account and safeguards –with its rights and obligations– the use of autonomous vehicles. Such proposal for legislation will be deeply based on those enacted bills mentioned in previous chapters. The purpose is to speed up drafting and passing of such legal frameworks to allow the use of autonomous vehicles.

Even this lack of legislation is –probably– the main problem of this innovative technology, this is not its only problem. Though the latter is really important and is hindering their testing, there are more issues regarding Autonomous. For this reason, the second part of the chapter is focused on solving such obstacles suggesting possible solutions, as well as risk spreading among the different stakeholders and its mitigation. The objective is, once again, to concentrate the efforts of all parties on pushing forward with autonomous technology. Finally, a guideline is presented to define and establish the basis of future “Autonomous-friendly” legislation and overcome all the outstanding challenges.

5.1 Proposal for Legislation

As mentioned above, the main objective of this section is the establishment of a new legal framework that allows the use of autonomous vehicles, with its consequent rights and obligations for those parties or stakeholders wishing to test and use such vehicles. This proposal should be a part of the Traffic Law or Highway Code, as appropriate. The need of legal framework responses to that, in many countries, this lack of legislation or legal vacuum is the main problem and obstacle to test and use autonomous vehicles in public roads. Nevertheless, and in my opinion, the main problem is that, on the one hand the Governments are not aware of the need for implementing new regulations by Autonomous manufacturers and developers, and on the other hand as a consequence of the legislative system and its reticence to change and slowness when drafting and approving new laws. In addition to all this, constant changes of government and new political parties could delay even more the approval of new legislative frameworks.

Besides, another problem is the divergence between countries and their enacted laws allowing the use of autonomous vehicles. But such divergence –as problem– is itself a solution for, on the one hand those countries without any bill and, on the other hand, for those that have drafted and approved several laws. The first countries, can adopt already enacted bills as a reference, in order to write their own laws. The latters for their part, with a mature legislative system, can see how their legislation is adopted by other countries as a reference, and in turn contribute to update their own laws, adding other articles proposed by the firsts. For this reason, and leveraging the momentum present in some legislations, arises the idea of establishing a common legal and regulatory framework for those countries wishing to review their laws, using enacted bills as a baseline. In addition, and as it will presented latter in this chapter, enacted laws are, in many cases, identical word for word. Then, adapting their content nationally, every country could have simply its own legislative framework.

Getting started, and as seen before, the United States has, if not the more, one of the most advanced legislative systems to allow the use of autonomous vehicles. In addition, several autonomous companies (e.g., Google, Tesla, Apple) have their headquarters in one of its States. As a consequence, they have pushed forward with local authorities to enact new laws to allow the use and test of such vehicles. Local authorities have realized that such companies have an important financial backing, and pursuing appropriate Autonomous policies can attract new stakeholders and investment. Its main consequence is that these States today are considered the reference in terms of passed legislation, with high “degrees of maturity”. In addition and thanks to some sources of information (previously mentioned) these laws are perfectly identified, which allows a quick and easy access to them. Such advantage is only found in the United States, due to in the rest of countries Autonomous regarding laws are not –in many cases– well defined and little more than an empty vessel. By contrast, those enacted by some States (e.g., Michigan, Florida, Nevada, North Dakota, Tennessee and District of Columbia) have been reviewed and even amended by other laws. In conclusion, the next step is going to be based on these enacted laws, to find if it is possible to establish a common legal framework for those countries wishing to allow the use of autonomous vehicles.

In accordance with the previous paragraph, the bills that have been taken as a frame of reference, and therefore as source, are those enacted in California, Michigan, Florida, Nevada, North Dakota, Tennessee and District of Columbia, described in Tab.3.1. The next step is a thorough analysis of the bills aforementioned to find, if possible, similarities between their different articles and structures. At first sight, several parallelisms are found due to the first article in almost all bills is the definition of those terms related with Autonomous, but the problem is that their structure changes significantly depending on which bill is considered. However, and after an in-depth study about what structure could better fit and properly reflect all the information, the Bill ND HB 1065 enacted by the State of North Dakota (from now on ND HB 1065) is considered as a reference. This one is, in my opinion, the best due to it shows effectively and in a simple way the content of its different articles. Such structure is the following:

01. Definitions.

02. Requirement for insurance or bond for testing autonomous vehicles.

03. Safety requirements for testing autonomous vehicles.

04. Requirements for registration for autonomous vehicles.

05. Manufacturer not liable for certain damages.

06. Department regulations to authorize operation of autonomous vehicles.

07. Endorsement to operate autonomous vehicles.

However, and as said before, such structure is going to be partially used as a reference due to it does not contain other important sections, provided in other bills. On the other hand, the order is also going to vary due to these changes. Besides, another article, stated by some bills (FL SB 52, NV SB 140 and MI SB 0169) is added to the legislation because of its relevance when discussing the use of autonomous vehicles. This article is the seventh and it is called *Prohibition of wireless communications devices*.

Finally, the proposed structure and its different articles are shown as follows:

1. *Definitions.*

2. *Requirement for insurance or bond for testing autonomous vehicles.*

3. *Safety requirements for testing autonomous vehicles.*

4. *Requirements for registration for autonomous vehicles.*

5. *Department regulations to authorize operation of autonomous vehicles.*

6. *Vehicle conversion and manufacturer's liability.*

7. *Prohibition of wireless communications devices.*

8. *Endorsement to operate autonomous vehicles.*

Once determined the structure of its different articles, the next step is going to be the definition of their content, once again, based on the different information available on the above defined sources.

Because of many of, or almost all the terms used to make reference to Autonomous vehicles are not regulated by law, the first step is the definition of such terms or expressions. This section is really important, and that is the reason why it appears in first position, due to if we want to refer to these terms, their meaning have to be defined by law. As seen above, the main problem of new technologies is that, as a consequence of their origins, laws do not regulate them, being in an alegal scenario. Such scenario, nearest to illegality in most cases, is complicating the progress of this innovative technology, meaning one of their biggest obstacles.

For this reason, the first objective is to find the terms that need to be defined, in order to lay the foundations of their legislation. With regard to Autonomous vehicles and according to already enacted laws, the main terms and expressions that need to be defined are: *autonomous technology, autonomous vehicle, operator (driver), manufacturer, autonomous mode, operate/operating and wireless communication device.*

Following, the previous terms are defined according to the definitions stated by the different legislations as a part of chapter ***1. Definitions***:

> *(1)(a) "**Autonomous technology**" means technology installed on a <u>motor vehicle</u> that has the capability to drive the motor vehicle without the active physical control or monitoring by a human operator.*
>
> *(b)The term excludes a motor vehicle enabled with active safety systems or driver assistance systems, including, without limitation, a system to provide electronic blind spot assistance, crash avoidance, emergency braking, parking assistance, adaptive cruise control, lane keep assistance, lane departure warning, or traffic jam and queuing assistant, unless any such system, alone or in combination with other systems, enables the vehicle on which the technology is installed to drive without the active control or monitoring by a human operator.*

The definition at point (1)(a) is based on TN SB 598, which agrees almost entirely with CA SB 1298, unless as above underlined, the latter considers *vehicle* as such. The rest of the bills, with definitions very similar, state such definition for *motor vehicles*, and that is the reason why the adjective *motor* is added to the previous definition. The section (b) is extracted from FL HB 1207, which along with NV SB 313 and ND HB 1065 state such exclusion. It is necessary to say that these three bills state such exclusion within the definition of *autonomous technology*. However, the rest of bills include the same subsection but as a part of *autonomous vehicle* definition. Due to this, and as shall be seen later, this exclusion has also been included as a subsection –with exactly the same content– to the definition of *autonomous vehicle*. The reason is that, although it may seem repetitive, this concept needs to be clear. Lastly, *autonomous technology* is also defined by MI SB 0169 but such definition contain the words *"..technology installed on a motor vehicle that has the capability to assist, make decisions for, or <u>replace an operator</u>"*. This could be a future problem due to –and as shown below– the operation of autonomous vehicles must be supervised at all times by an *operator*. In addition, this bill also uses the adjective *automated* in reference to both technology and vehicle, but in order to follow a common framework, this term is going to be replaced by *autonomous* (the use of the latter is currently widespread).

The following definition makes reference to the term *autonomous vehicle*, and is shown below:

> *(2)(a) "**Autonomous vehicle**" means a motor vehicle that is equipped with autonomous technology.*
>
> *(b) The term excludes a motor vehicle enabled with active safety systems or driver assistance systems, including a system to provide electronic blind spot assistance, crash avoidance, emergency braking, parking assistance, adaptive cruise control, lane keep assistance, lane departure warning, or traffic jam and queuing assistance, unless a system, alone or in combination with other systems, enables the vehicle on which the technology is installed to drive without active control or monitoring by a human operator.*

As in the previous definition, the section (2)(a) is based on ND HB 1065, broadly similar to the definitions stated by other bills. Concerning to section (b), and as previously mentioned, the paragraph extracted from DC B19-0931 –also stated by MI SB 0169– is once again added to the legislation. As seen above, bills consider such subsection in one or another definition, but in such proposal it is added to both definitions to increase its significance. Finally and in case that the proposed exclusion might seem repetitive due to the fact that it appears in two different definitions, it should be included only in the first definition (according to its preference). For this reason this last definition is open to change.

The following and third definition refers to *operator* and is shown below:

> *(3) An "**operator**" of an autonomous vehicle is the person, with <u>a valid driver's license</u>, who does either of the following:*
>
> *(a) Operates an autonomous vehicle.*
>
> *(b) Is seated in the driver's seat.*
>
> *(c) If there is no person in the driver's seat, causes the autonomous technology to engage.*

This third definition is obtained from the combination of CA SB 1298 and MI SB 0169. The content proposed by both bills is slightly incomplete, but such lack is solved by their combination. In addition, only these two bills define the term *operator*, despite its relevance. On the other hand, another similar definition is shown in DC B19-0931, but it is referred to *driver* as *human operator of a motor vehicle with a <u>valid driver's license</u>*. This has also been added to the proposed definition to emphasize the obligatory nature of an operator with a valid driver's license. Besides and as will be discussed later, the possession of a valid driver´s license by those parties desiring to test such vehicles, is going to be considered as a necessary requirement to such purpose. Thus this definition already establishes interrelationships with coming articles.

The fourth term makes reference to *manufacturer´s* party. This definition is also needed because of its common use in several articles and due to it is key during the distribution of responsibilities between parties in case of failure. The term *manufacturer* refers to:

> *(4) A "**manufacturer**" of autonomous technology is the person that originally manufactures a <u>motor</u> vehicle and equips autonomous technology on the originally completed vehicle or, in the case of a <u>motor</u> vehicle not originally equipped with autonomous technology by the vehicle manufacturer, the person that modifies the vehicle by installing autonomous technology to convert it to an autonomous vehicle after the <u>motor</u> vehicle was originally manufactured.*

This term is only defined by three bills, being the one present in CA SB 1298 considered as the best. The other two bills state definitions that are not as accurate as this one. For example NV AB 511 only mentions *original manufacturer*, so in case a motor vehicle had been modified by a second manufacturer, the latter would not be considered as manufacturer and that could be another problem. Finally the adjective motor is added to the definition in line with other definitions.

The next definition concerns the verb *operate* and its respective gerund *operating*. The definition of these words is again required when we talk about autonomous vehicles. For this reason, both terms are defined as follows:

> *(5) "**Operate**" or "**operating**" means one or more of the following:*
>
> *(a) Being in actual physical control of a vehicle.*
>
> *(b) Causing an automated motor vehicle to move under its own power in automatic mode upon a highway or street ~~regardless of whether the person is physically present in that automated motor vehicle at that time~~.*

As opposed to the other definitions, this one only appears in MI SB 0169. The only difference is, and as seen above, the fact that an operator is not needed inside the autonomous vehicle during its operation. This sentence is deleted because, as seen below, an operator should be inside the vehicle at all times to lawfully drive the vehicle through any public road. As in previous cases, the intention of such proposal is to maintain the same criterion in all this guideline.

One of the last terms defined is *automatic mode* due to, even it is not vastly used, it tends to appear when we speak about autonomous vehicles. For this reason, it is defined as follows:

> *6) "**Autonomous mode**" means the mode of operating an ~~automated motor~~ autonomous vehicle when ~~automated~~ autonomous technology is engaged to enable the motor vehicle to operate without any control or monitoring by an operator.*

This definition only appears in MI SB 0169, which as stated above, uses the adjective *automated* instead of *autonomous*. For this reason, and to be consistent with other definitions, that is the only change.

Finally, and as a consequence of a future article, the definition of the expression *wireless communication device* is added to this section. An entire article is based on this device and none legislation takes its meaning into account. Is therefore established:

> *(7) "**Wireless communications device**" means any handheld device used or capable of being used in a handheld manner, that is designed or intended to receive or transmit text or character-based messages, access or store data, or connect to the Internet or any communications service and that allows text communications.*

Even though some bills state legislation about such devices (e.g., FL SB 52, NV SB 140 and MI SB 0169), the only one that defines its meaning is FL SB 52 and it is shown above. Thus and as said before, its definition has been considered necessary despite it is not usually covered by enacted legislation.

Once defined all key concepts, which are going to constantly appear in coming articles, the next step is based on the demanded requirements for those parties wishing to test autonomous vehicles on public roads. The first requirement is related to insurance, which according to the literature is written as follows:

> *2. Requirement for insurance or bond for testing autonomous vehicles.*
>
> *Prior to the start of testing ~~in this State~~ on a public road within this Country, the manufacturer performing the testing shall obtain an instrument of insurance, surety bond, or proof of self-insurance in the amount of five million dollars ($5,000,000), and shall provide evidence of the insurance, surety bond, or self-insurance to the Department.*

The original article appears in CA SB 1298, being the only change the words *on a public road within this country* that replace the terms *in this state*. Due to their state nature, their scope of application is also state, as laid down above. For this reason, and with the objective of increasing such scope of application nationally, the words *in this state* are replaced by *within this country*. The terms *on a public road* are also added to establish where they can be tested (including highways and streets). On the other hand, the word *Department* means the national department responsible for road traffic. Once explained the content of this article, it is important to discuss two concepts. The first one is the set amount for the instrument of insurance in five million dollars. This threshold is the same in all bills that contain such requirement (CA SB 1298, FL HB 1207 and ND HB 1065) and due to this it is maintained in this proposal. The second one is that, when such trials take place more usual, this instrument of insurance could be another challenge or obstacle for manufacturers because of its high cost. Then, as the number of trials growsm, this amount should be reduced in the same way. The aim of this proposal is not other than stimulate and encourage manufacturers to test, but without forgetting in any way whatsoever any safety requirement. Another important thing is that every county should adapt this amount to its own economy and local currency. Finally, this instrument of insurance can be seen as an opportunity to insurance companies to enter the business. It is necessary to keep in mind that they have to be an important player in this business and give an extra momentum to autonomous vehicles.

The second requirement, established with the aim of regulating the use of autonomous vehicles refers to safety as follows:

> *3. Safety requirements for testing autonomous vehicles.*
>
> *If an autonomous vehicle is being tested on a ~~highway~~ <u>public road</u> within ~~this State~~ <u>this Country</u>, an individual who is the operator must be:*
>
> *(1) An employee, contractor, or other person designated or otherwise authorized by that manufacturer of ~~automated~~ <u>autonomous</u> technology.*
>
> *(2) Seated in a position which allows the individual to take immediate manual control of the autonomous vehicle.*
>
> *(3) Monitoring the safe operation of the autonomous vehicle; and*
>
> *(4) Capable of taking over immediate manual control of the autonomous vehicle in the event of a failure of autonomous technology or other emergency.*

The points (2), (3) and (4) of this article are based on the requirements established by ND HB 1065. In order to complete its significance, the requirement at the first point (1) is added to the article. This point appears on MI SB 0169 and its purpose is the limitation of the term operator only to certain persons related with the manufacturer. On the other hand, another purpose of this point is to hold responsible the manufacturer in case of accident, and hence, increasing their liability. Other changes are based on maintaining the lexicon proposed in previous articles.

The following article of this proposal refers to the requirements demanded for registration of autonomous vehicles. Such requirements are shown as follows:

4. Requirements for registration for autonomous vehicles.

(1) An autonomous vehicle may not be registered in this ~~State~~ <u>Country</u> unless the autonomous vehicle meets all ~~federal~~ <u>national</u> standards and regulations that are applicable to a motor vehicle.

(2) An autonomous vehicle may not be tested or operated on a ~~highway~~ <u>public road</u> within this ~~State~~ <u>Country</u> unless the autonomous vehicle is:

(a) Equipped with a means to engage and disengage the autonomous technology which is easily accessible to the individual who is the operator of the autonomous vehicle.

(b) Equipped with a visual indicator located inside the autonomous vehicle which indicates when autonomous technology is operating the autonomous vehicle.

(c) Equipped with a system to safely alert the operator if an autonomous technology failure is detected while the autonomous technology is engaged, and when an alert is given, the system shall do either of the following:

(i) Require the operator to take control of the autonomous vehicle.

(ii) If the operator does not or is unable to take control of the autonomous vehicle, the autonomous vehicle shall be capable of coming to a complete stop.

(d) Capable of allowing the operator to take control in multiple manners, including, without limitation, through the use of the brake, the accelerator pedal, or the steering wheel, and it shall alert the operator that the autonomous technology has been disengaged.

(e) Equipped with a separate mechanism, in addition to, and separate from, any other mechanism required by law, to capture and store the autonomous technology sensor data for at least 30 seconds before a collision occurs between the autonomous vehicle and another vehicle, object, or natural person while the vehicle is operating in autonomous mode. The autonomous technology sensor data shall be captured and stored in a read-only format by the mechanism so that the data is retained until extracted from the mechanism by an external device capable of downloading and storing the data. The data shall be preserved for three years after the date of the collision.

(f) Capable of being operated in compliance with applicable motor vehicle laws and traffic laws of this ~~State~~ Country.

The content of this fourth point is based on a combination of two different bills. The first point (1) and second (2) in subsections (a), (b) and (f) appear on ND HB 1065, and the rest of the article was extracted from CA SB 1298. The point (c) of both bills has a similar content, but in case of CA SB 1298, an assumption not considered in ND HB 1065 is presented. Such assumption considers a scenario where the operator *is unable to take control of the autonomous vehicle*. In addition this could be seen as a consideration, although it is not cited as such, of a scenario in which the operator is not physically inside the autonomous vehicle, being then a vehicle "entirely autonomous". In addition, such bill considers in several and different articles such hypothesis (e.g*., "… are necessary to ensure the safe operation of autonomous vehicles on public roads, with or without the presence of a driver inside the vehicle*; or "… *the department shall hold public hearings on the adoption of any regulation applicable to the operation of an autonomous vehicle without the presence of a driver inside the vehicle")*, which has never seen before in any other bill. The points (c), (d) and (e) are based on respective points (C), (D) and (G) from CA SB 1298. Finally, in (e) another new concept is presented and has a similar meaning to black boxes in civil aviation. Even thought it only appears in this bill, its content is really important due to in case of accident it could ease to establish the causes of the accident.

The next article of this proposal for legislation makes reference to regulations that need to be adopted by the corresponding Department to authorize the use and operation of such vehicles. The regulations of this fifth article are read as follows:

5. Department regulations to authorize operation of autonomous vehicles.

(1) The Department shall adopt regulations authorizing the operation of autonomous vehicles on ~~highways~~ public roads within this ~~State~~ Country.

(2) The regulations required to be adopted by subsection (1) must:

(a) Set forth requirements that an autonomous vehicle must meet before it may be operated on a ~~highway~~ public road within this ~~State~~ Country.

(b) Set forth requirements for the insurance that is required to test or operate an autonomous vehicle on a ~~highway~~ public road within this ~~State~~ Country.

(c) Establish minimum safety standards for autonomous vehicles and their operation.

(d) Provide for the testing of autonomous vehicles.

(e) Restrict the testing of autonomous vehicles to specified geographic areas; and

(f) Set forth such other requirements as the Department determines to be necessary.

Such article is entirely based on ND HB 1065 and as in previous articles the changes are only lexical, on the one hand to adapt its scope nationally and on the other hand to adapt to every public road –either highway or street–.

Another important point of this proposal for legislation is liability of manufacturer in case of failure, but when such failure is produced by a third party different than the original manufacturer. Therefore, and due to manufacturer's liability is already high enough for this vehicles so technologically dependants, an exclusion –or limitation– of liability is added to this proposal. Nevertheless, this clause will be effective provided that one of the following conditions is satisfied:

6. Vehicle conversion and manufacturer's liability.

(1) The <u>original</u> manufacturer of a <u>motor</u> vehicle is not liable and shall be dismissed from any action for alleged damages resulting from any of the following unless the defect from which the damages resulted was present in the vehicle when it was manufactured:

(a) The conversion or attempted conversion of the vehicle into an ~~automated motor~~ <u>autonomous</u> vehicle by ~~another person~~ <u>a third party</u>.

(b) The installation of equipment in the vehicle by ~~another person~~ <u>a third party</u> to convert it into an ~~automated motor~~ <u>autonomous</u> vehicle.

(c) The modification by ~~another person~~ <u>a third party</u> of equipment that was installed by the <u>original</u> manufacturer in an ~~automated motor~~ <u>autonomous</u> vehicle specifically for using the vehicle in automatic mode.

(2) A subcomponent system producer is not liable in a product liability action for damages resulting from the modification of equipment installed by the subcomponent system producer to convert a <u>motor</u> vehicle to an ~~automated~~ <u>autonomous</u> vehicle unless the defect from which the damages resulted was present in the equipment when it was installed by the subcomponent system producer.

In this case the liability, or better said, its exclusion by the original manufacturer is taken into account by several bills (DC B19-0931, FL HB 1207, MI SB 0169, MI SB 0663, NV SB 313 and ND HB 1065) but the above article is based entirely on MI SB 0663. Its changes are mainly two: the first ones are those to maintain the lexicon, and the second one is that the expression *another person* is replaced by *a third party* and the adjective *original* is added as a complement to the noun *manufacturer*. The first change is made because all bills –with the exception of mentioned MI SB 0663– use such expression and it has a wide meaning. The second one establishes who is the *original manufacturer* in order to define which party is exempt of liability. In addition, as in the previous case, several bills refer to such terms when they talk about manufacturer's liability.

One of the last articles refers to prohibition of wireless communication devices, but limiting their application to motor vehicles that are not fully automated or Autonomous. This law only appears in few bills like FL SB 52, NV SB 140 and MI SB 0169, but its application in test is certainly very important. Due to multiple parties are involved when testing, the communication between them is basically through such wireless devices. Nevertheless, without this clause or exclusion of application the use of these devices will be prohibited and punished. Thus, and keeping in mind the purpose of easing the test and use of autonomous vehicles, the article is shown as follows:

> *7. Prohibition of wireless communications devices.*
>
> *(1) A person shall not, while operating a motor vehicle:*
>
> *(a) Manually type or enter text into a cellular telephone or other handheld wireless communications device, or send or read data using any such device to access or search the Internet or to engage in nonvoice communications with another person, including, without limitation, texting, electronic messaging and instant messaging.*
>
> *(b) Use a cellular telephone or other handheld wireless communications device to engage in voice communications with another person, unless the device is used with an accessory which allows the person to communicate without using his or her hands, other than to activate, deactivate or initiate a feature or function on the device.*
>
> *(2) Paragraph (1) does not apply to a motor vehicle operator who is:*
>
> *(a) Operating an autonomous vehicle in autonomous mode.*

This article is entirely based on NV SB 140, and its most important section is the point (2)(a) due to it does not apply to an *autonomous vehicle in autonomous mode*. Besides, this point only applies if the autonomous vehicle is operated in autonomous mode, because if not, it would be considered as a normal motor vehicle and the use of these devices is totally prohibited. Lastly, this article is the main reason why the expression *wireless communication device* was defined in the first article.

The eighth and last article of the proposal for legislation, which does not have so much relevance as the others, endorses the use of autonomous vehicles without the active supervision of an operator. This article is written as follows:

8. Endorsement to operate autonomous vehicles.

The Department shall by regulation establish a driver's license endorsement for the operation of an autonomous vehicle on ~~the highways~~ public roads of this ~~State~~ Country. The driver's license endorsement described in this section must, in its restrictions or lack thereof, recognize the fact that an individual is not required to actively drive an autonomous vehicle.

The content described by both NV AB 511 and ND HB 1065 is exactly the same word for word, so the only needed changes are those to establish their scope nationally.

This proposal for legislation, as said before, is thought to be a query tool for either those countries that have not already enacted any law about autonomous vehicles and their use, or those that, having enacted laws yet, want to review or update their content. This proposal should also be constantly reviewed, preventing it from being outdated, according to Autonomous' development. This is strictly necessary due to, on the one hand technological development has always a faster velocity than legislation, and on the other hand this proposal needs to be dynamic and open to change. This will avoid future problems, typical of other legislations (e.g., reticence to chance, rigidity) and contribute to ease the implementation of autonomous vehicles.

In conclusion, such guideline can be very useful for those parties wishing to allow the use of autonomous vehicles in their public roads, but at the same time under a legal framework that regulates their rights and obligations. It is also a temporary guideline due to it needs to be constantly reviewed and updated. Besides it has to be a part of the Highway Code, with its required changes to be consistent with the latter, but without forgetting that some changes in the Highway Code are going to be needed to allow the use of Autonomous.

Below and attached is the "**Proposal for Legislation of Autonomous Vehicles**" in its entirety.

1. Definitions

(1)(a) "Autonomous technology" means technology installed on a motor vehicle that has the capability to drive the motor vehicle without the active physical control or monitoring by a human operator.

(b)The term excludes a motor vehicle enabled with active safety systems or driver assistance systems, including, without limitation, a system to provide electronic blind

spot assistance, crash avoidance, emergency braking, parking assistance, adaptive cruise control, lane keep assistance, lane departure warning, or traffic jam and queuing assistant, unless any such system, alone or in combination with other systems, enables the vehicle on which the technology is installed to drive without the active control or monitoring by a human operator.

(2)(a) "Autonomous vehicle" means a motor vehicle that is equipped with autonomous technology.

(b) The term excludes a motor vehicle enabled with active safety systems or driver assistance systems, including a system to provide electronic blind spot assistance, crash avoidance, emergency braking, parking assistance, adaptive cruise control, lane keep assistance, lane departure warning, or traffic jam and queuing assistance, unless a system, alone or in combination with other systems, enables the vehicle on which the technology is installed to drive without active control or monitoring by a human operator.

(3) An “operator” of an autonomous vehicle is the person, with a valid driver’s license, who does either of the following:

(a) Operates an autonomous vehicle.

(b) Is seated in the driver’s seat.

(c) If there is no person in the driver’s seat, causes the autonomous technology to engage.

(4) A “manufacturer” of autonomous technology is the person that originally manufactures a motor vehicle and equips autonomous technology on the originally completed vehicle or, in the case of a motor vehicle not originally equipped with autonomous technology by the vehicle manufacturer, the person that modifies the vehicle by installing autonomous technology to convert it to an autonomous vehicle after the motor vehicle was originally manufactured.

(5) “Operate” or “operating” means one or more of the following:

(a) Being in actual physical control of a vehicle.

(b) Causing an automated motor vehicle to move under its own power in automatic mode upon a highway or street.

(6) “Autonomous mode” means the mode of operating an autonomous vehicle when autonomous technology is engaged to enable the motor vehicle to operate without any control or monitoring by an operator.

(7) “Wireless communications device” means any handheld device used or capable of being used in a handheld manner, that is designed or intended to receive or transmit text or character-based messages, access or store data, or connect to the Internet or any communications service and that allows text communications.

2. Requirement for insurance or bond for testing autonomous vehicles.

Prior to the start of testing on a public road within this Country, the manufacturer performing the testing shall obtain an instrument of insurance, surety bond, or proof of self-insurance in the amount of five million dollars ($5,000,000), and shall provide evidence of the insurance, surety bond, or self-insurance to the Department.

3. Safety requirements for testing autonomous vehicles.

If an autonomous vehicle is being tested on a public road within this Country, an individual who is the operator must be:

(1) An employee, contractor, or other person designated or otherwise authorized by that manufacturer of autonomous technology.

(2) Seated in a position which allows the individual to take immediate manual control of the autonomous vehicle.

(3) Monitoring the safe operation of the autonomous vehicle; and

(4) Capable of taking over immediate manual control of the autonomous vehicle in the event of a failure of autonomous technology or other emergency.

4. Requirements for registration for autonomous vehicles.

(1) An autonomous vehicle may not be registered in this Country unless the autonomous vehicle meets all national standards and regulations that are applicable to a motor vehicle.

(2) An autonomous vehicle may not be tested or operated on a public road within this Country unless the autonomous vehicle is:

(a) Equipped with a means to engage and disengage the autonomous technology which is easily accessible to the individual who is the operator of the autonomous

vehicle.

(b) Equipped with a visual indicator located inside the autonomous vehicle which indicates when autonomous technology is operating the autonomous vehicle.

(c) Equipped with a system to safely alert the operator if an autonomous technology failure is detected while the autonomous technology is engaged, and when an alert is given, the system shall do either of the following:

(i) Require the operator to take control of the autonomous vehicle.

(ii) If the operator does not or is unable to take control of the autonomous vehicle, the autonomous vehicle shall be capable of coming to a complete stop.

(d) Capable of allowing the operator to take control in multiple manners, including, without limitation, through the use of the brake, the accelerator pedal, or the steering wheel, and it shall alert the operator that the autonomous technology has been disengaged.

(e) Equipped with a separate mechanism, in addition to, and separate from, any other mechanism required by law, to capture and store the autonomous technology sensor data for at least 30 seconds before a collision occurs between the autonomous vehicle and another vehicle, object, or natural person while the vehicle is operating in autonomous mode. The autonomous technology sensor data shall be captured and stored in a read-only format by the mechanism so that the data is retained until extracted from the mechanism by an external device capable of downloading and storing the data. The data shall be preserved for three years after the date of the collision.

(f) Capable of being operated in compliance with applicable motor vehicle laws and traffic laws of this Country.

5. Department regulations to authorize operation of autonomous vehicles.

(1) The Department shall adopt regulations authorizing the operation of autonomous vehicles on public roads within this ~~S~~ Country.

(2) The regulations required to be adopted by subsection (1) must:

(a) Set forth requirements that an autonomous vehicle must meet before it may be operated on a public road within this Country.

(b) Set forth requirements for the insurance that is required to test or operate an

autonomous vehicle on a public road within this Country.

(c) Establish minimum safety standards for autonomous vehicles and their operation.

(d) Provide for the testing of autonomous vehicles.

(e) Restrict the testing of autonomous vehicles to specified geographic areas; and

(f) Set forth such other requirements as the Department determines to be necessary.

6. Vehicle conversion and manufacturer's liability.

(1) The original manufacturer of a motor vehicle is not liable and shall be dismissed from any action for alleged damages resulting from any of the following unless the defect from which the damages resulted was present in the vehicle when it was manufactured:

(a) The conversion or attempted conversion of the vehicle into an autonomous vehicle by a third party.

(b) The installation of equipment in the vehicle by a third party to convert it into an autonomous vehicle.

(c) The modification by a third party of equipment that was installed by the original manufacturer in an autonomous vehicle specifically for using the vehicle in automatic mode.

(2) A subcomponent system producer is not liable in a product liability action for damages resulting from the modification of equipment installed by the subcomponent system producer to convert a motor vehicle to an autonomous vehicle unless the defect from which the damages resulted was present in the equipment when it was installed by the subcomponent system producer.

7. Prohibition of wireless communications devices.

(1) A person shall not, while operating a motor vehicle:

(a) Manually type or enter text into a cellular telephone or other handheld wireless communications device, or send or read data using any such device to access or search the Internet or to engage in nonvoice communications with another person, including, without limitation, texting, electronic messaging and instant messaging.

(b) Use a cellular telephone or other handheld wireless communications device to engage in voice communications with another person, unless the device is used with an accessory which allows the person to communicate without using his or her hands, other than to activate, deactivate or initiate a feature or function on the device.

(2) Paragraph (1) does not apply to a motor vehicle operator who is:

(a) Operating an autonomous vehicle in autonomous mode.

8. Endorsement to operate autonomous vehicles.

The Department shall by regulation establish a driver's license endorsement for the operation of an autonomous vehicle on public roads of this Country. The driver's license endorsement described in this section must, in its restrictions or lack thereof, recognize the fact that an individual is not required to actively drive an autonomous vehicle.

5.2 Proposal for Improvements

The previous section described a proposal for legislation useful for those countries wishing to add *Autonomous-friendly* laws to their road legislations. However, despite legislative problem is the most significant, it is not the unique against which autonomous vehicles have to deal. For all these reasons, the objective of this section is to provide another guideline to solve such problems, establishing possible solutions, keeping in mind the objective of helping to facilitate the use of autonomous vehicles.

As has been said so many times, the main problem is the legal vacuum in terms of enacted legislation to allow the use of Autonomous. However, research –and consequently autonomous technologies– is constitutionally considered, but the problems arise when such technology is tested. However, this is a typical problem when we speak about new technologies that are in advanced stages of development, and the next step is their testing, which are not legally allowed and, consequently, are considered illegal. But this should not be that way. Although the legal framework is really rigid and not open to change, there are some constitutional mechanisms that allow such change. This means that if an article needs to be amended, the Government has enough tools to do that, and that is what really matters. If Governments do not take a proactive attitude towards new technologies and persist in ignoring them, such changes cannot be done. For this reason, it is their duty and obligation to provide new regulatory frameworks to allow the use of autonomous vehicles, provided that such technologies are at advanced stages of development and under strict monitoring. Ultimately, every country is able to enact laws, as a new part or complement of their Road

Traffic Law, to allow the use of autonomous vehicles. On the other hand, having a proposal for legislation based on already enacted bills, can be really useful for some countries, which only had to adjust its articles to their national requirements –if necessary–, shortening procedures and periods. Another important point is that the Proposal for Legislation is written under an Anglo-Saxon legislative system, also called Common Law. Therefore and if necessary, every country should adapt it for its own system (e.g., Civil Law, Muslim Law, Customary Law or Mixed System). Nevertheless this is a minor problem.

With regard to technological problem, more than a problem is a challenge. This could have been a problem years ago, but today research, investment and development of this technology are in highpoint. However, its main challenge is to overcome those problems inherent in any new technology. In addition, the fact that the autonomous technology is going to be the driver or which substitutes the person who drives the vehicle, and consequently partially responsible in case of accident, contributes to put even more pressure on this technology. It should also be noted that in many cases, manufacturers will have to face liability, and therefore also strict quality and security standards imposed by government regulation. This means, unequivocally, an additional provision of resources on the part of manufacturers, increasing its cost of production affecting its profitability. The objective is finding the optimum technology that never fails, but as we all know, technological risk is ever present and a situation of no risk can never be obtained. For this reason, the main responsible party and holder of such risk has to be the manufacturer. However, we should not forget that, if our ultimate goal –as a society– is the use of autonomous vehicles, both their users and all concerned parties will be –to some extent– exposed to such risk. But finally, that should not be a great problem due to if there is something that makes autonomous technology remarkable is its capacity to improve our quality of life and life expectancy.

Security and liability are another two important factors. As was mentioned in the previous paragraph, autonomous vehicles are designed to improve both quality of life and life expectancy of the population. But, in spite of this, such vehicles are going to be unavoidably involved in accidents, including some scenarios resulting in death. The concept of no risk for such technology is idyllic but false due to it can never be obtained. Therefore, and even autonomous technology is thought to reduce the number of road accidents and fatalities, some of them will be caused by Autonomous. For this reason and returning to security and liability matters, if human driver's role is going to be replaced by autonomous technology, the first is going to base their security on the latter and consequently this technology must surpass strict security standards. Due to all of these reasons, developers and manufacturers have to design technology with a high degree of reliability and totally tested, to ensure that in case of failure mechanism are in place to solve such incidence guaranteeing maximum safety for both occupants of the vehicle and for the rest of drivers and users of public roads and surroundings. In case of fatalities, because of a flagrant manufacturer failure, such party may have to face civil and criminal liability. Therefore, the responsibility of the manufacturer to fulfil with security standards and quality controls could contribute to avoid worse consequences. On the other hand, it is necessary to define those cases in which manufacturer is not responsible for damages. Thus, it is protected in case of a third party had modified the autonomous vehicle, and as a consequence of this the failure occurred. For this reason the article 6 of the Proposal for Legislation (*Vehicle conversion and manufacturer's liability*) was taken into account and added to the guideline. This article is base on excluding original manufacturer's liability when the vehicle has been modified by a third party different than the original. In this way, original manufacturer disclaims all liability or responsibility for acts and omissions of third parties. This is really important due to, so far, risk spreading had been largely based on manufacturer (also considering developers) practically as a unique holder. Such distribution was not well balanced because of the lack of rights on the part of the manufacturer, and considering that the problem is caused by another party, it is fair that such party faces liability instead of manufacturer. Finally, manufacturer is going to have, always, the majority of risks due to it is the most interested party in autonomous technology and as a consequence (giving rise to criteria for assessing the risk-return trade-off) its profits are also going to be great.

Another problem of Autonomous has its origins in the large number of sensors equipped in the vehicle. Such sensors are installed to obtain a lot of information from it surrounding (e.g., traffic and weather conditions, accidents, routes) and, through a data processing, they will allow the autonomous robot to drive the vehicle according to those generated scenarios and taking its own decisions. Besides, the use of these sensors will establish communication with other vehicles, information points or servers, with which is going to share multiple and numerous information. As everybody knows, all the information generated nowadays contains personal data coming from the party that sent it out. Consequently, and due to the enormous amount of information generated, inarguably it will contain personal characteristics and problems could arise due to its processing and –if it were the case– its use. Some examples of the aforementioned are requests for route to on board-computers or times of operation. Such information, if it is not managed according to strict standards defined by the Data Protection Law, may turn into a mayor problem affecting the use of these vehicles. Despite a lot of countries have enacted their own Data Protection Laws, it is likely that they are going to need a deep revision to adapt their content to autonomous vehicles. Besides, if the vehicle stores such information, the manufacturer could probably easily access to it, and consequently use it for other purposes than the original, based on the supervision of the vehicle. For this reason and in connection with the previous, Data Protection Laws need to reinforce Autonomous user rights by penalising the inappropriate use of such information by third parties. Another possible and simpler solution could be the limitation of the storage capacity of such devices, not allowing them to have access to second storage devices, for example thorough Internet. Other solution could be limiting to manufacturers their time to store information, moment from which they should remove all the information.

As seen in the previous paragraph, the technology used by autonomous vehicles is going to be able to establish communication with other vehicles, servers or external agents. In addition, they will also have instant access to Internet. For this reason and due to their exposure to numerous external agents or malicious software, cyber security arises as another possible problem of this technology. In addition, such external agents could have access to the information generated or stored by the vehicle, as seen in the previous paragraph. But this is not its biggest problem due to exist one even more important. This one is hackers and their cyber attacks. Already exists a video in which a hacker, only using a laptop, starts an automated vehicle and drives it throughout a parking. For this reason, in case the installed antivirus and antispyware were not enough, external agents could operate an Autonomous with its consequences. In addition, such vehicles could be seen as new tool to commit crimes only using a laptop and reducing their exposition. For this reason, and aware of the fact that hackers are always one step ahead of developers, taking advantage of such failures –always– existing in software, the likelihood of this event is noteworthy. Thus, and in case of failure, manufacturers have to install a physical brake to keep the vehicle stationary. Besides, such alternative brake as to be physical, as mentioned above, avoiding other electronic devices that could be easily deactivated. A clear example of this is the parking brake. Old manual parking brakes kept the vehicle stationary regardless of anyone trying to move it. However, today such devices have been replaced by electronic systems, which could be easily manipulated and deactivated by an extern agent, due to their electronic features. But this hypothesis only considers parked vehicles, so the next step is what would happen if the vehicle was in autonomous mode. In this scenario, developers and manufacturers have to design, again, additional mechanisms and methods to face such intrusions. On the one hand, trying to ensure that such agents could not have access to the vehicle, using more restrictive firewalls and, in case they had managed to enter to the autonomous system, the vehicle had to be able to come to stop. The latter is already included in the proposed legislation, in its article 4 section (c)(ii), which states that "*if an autonomous technology failure is detected while the autonomous technology is engaged and the operator does not or is unable to take control of the autonomous vehicle, the autonomous vehicle shall be capable of coming to a complete stop"*. In this way the effect of such malicious software would be reduced in order to minimize its consequences. Finally, this last point is really important due to, as already said, it would reduce and limit the effectiveness of a cyber attack, and consequently would help to consolidate the reliability of this technology, so exposed to external cyber risks.

The last noteworthy problem of autonomous vehicles is ethics and algorithms of death. According to ethics, it should be realised that every new technology will have to deal with such factor. In addition to this, in those technologies that could cause human accidents, including victims in some cases, this factor will raise even more. This is why, and provided that the degree of reliability of such technology is high enough, it should not represent a major problem, being easy to solve. An example of mentioned above is a lift. The first time this technology was introduced, it was used to transport persons vertically, to a selected floor, in an autonomous way. As can be seen, the similarities between both technologies are really close. At that time, ethic factors of such technology were numerous, but thanks to its reliability and passenger safety, together with its high social impact and revolution, ethic did not represent a major problem and its use become quickly widespread and entirely normal. Besides, have been accidents with deaths as a result of technological failures of a lift, and liability and risk-sharing mechanisms have functioned properly to date. On the other hand, those algorithms of death, which should be necessary implemented in autonomous vehicles in order to decide which path –in case of accident– has the lowest damage or likelihood of collision, can be another problem. However, and as in the previous case, this should not represent a major problem due to, in the worst case scenario, the vehicle will maintain its path trying to reduce risk to third parties. Consequently, in case of death, these will be owing to the recklessness of the victim. In addition, such algorithms should consider all possible scenarios and minimize, in case of accident, its results. In conclusion, the function of these algorithms in not based on deciding who dies and who not, as seen in previous chapters, but choosing the scenario with the lowest damage or likelihood of collision. Nevertheless it is the job of developers to take into account all the possible cases, through computational models, and propose solutions to the vehicle enabling it to become “autonomous” taking decisions. In addition, we should not forget that its operator or driver could take it own decision, being then responsible for the act.

In conclusion, autonomous vehicles are going to face multiple and varied problems, but we should not forget that this is totally normal for an innovative technology, which has never seen before and in advanced stages of development. Besides, it is irrefutable that its future benefits are going to be infinitively greater than its problems. For this reason, and due to such benefit will have an effect on society as a whole, all the stakeholders (e.g., manufacturers, public authorities and associations, drivers and pedestrians) have to join forces to solve such problems in the best way. Although manufacturers will be the more benefited party, and consequently also the main risk holder, public authorities will also gain advantage. For this reason, the firsts should encourage the latters to modify their current legislations in order to ease the use of autonomous vehicles. Without an active role of manufacturers this goal will never be reachable. Consequently, their technological proposals have to be reliable and rigorous in order to be well considered by Countries, as new sources of development, and then having to review their current legal frameworks to allow their use on public roads. In conclusion, if all stakeholders are concerned and do their part according to their know-how, it is irrefutable that this technology will be realised on public roads in short- to mid-term, with its benefits to society.

6 Conclusion

Currently, the use of new technologies that allow a vehicle to drive or operate with a high degree of automation (e.g., cruise control, lane keeping assist system) has already proven its reliability and security. In addition the use of such devices has contributed to reduce both fuel consumption and traffic fatalities. Thus, it is demonstrated that technologies with high degrees of automation are more efficient and secure than traditional ones. However, autonomous vehicles are thought to go even further, being not necessary the human driver as it is conceived today. In addition, in many cases such element will not be necessary due to the autonomous vehicle is going to be able to take its own decisions, driving the vehicle from a given certain place to another.

Today, this new driving concept is a reality, and several tests have been already performed in public roads, but subject to several restrictions. In addition, many laws do not consider the use of autonomous vehicles, viewed as illegal by many countries. This results in the cliché that "*the law always lags behind the development of new technologies*". But such affirmation is not true due to law exist even before than any given technology. The problem is its interpretation, and consequently any act that is not considered by the law is then categorised as illegal. But if autonomous vehicles want to become a reality, legal frameworks need to be reviewed. For example, in some cases it would be solved by extending the interpretation of existing laws. However, in other cases far reaching changes are needed. For this reason a Proposal for Legislation is presented in this paper. This can be used, by those countries wishing to allow the use of autonomous vehicles, as a reference tool and draft to write up their own laws or legal frameworks.

However, despite legislative problem is the most significant, there are other important problems that those parties interested in autonomous vehicles will have to face. But as in the previous case, the solution –despite it is not easy– is really similar. Every party has to take a proactive role solving those problems within reach, using their expertise to solve them. Even manufacturer is –and has to be– the main risk-holder, other parties have to take part to boost the technological development of these vehicles. In conclusion, autonomous vehicles are going to face multiple and varied problems, but we should not forget that this is totally normal for an innovative technology, which has never seen before and in advanced stages of development. Besides, it is irrefutable that its future benefits are going to be infinitively greater than its problems. For this reason, and due to such benefit will have an effect on society as a whole, all the stakeholders, using their know-how, have to join forces to solve such problems in the best way.

List of References

SMITH B. W. [2014]: Automated Vehicles Are Probably Legal in the United States. 1 Tex. A& M L. Rev. 411 (2014).

DOKIC J.; MÜLLER B.; MEYER G. [2015]: European Roadmap Smart Systems for Automated Driving. Berlin, April 1st, 2015.

HILGENDORF E. [2015]: AUTONOMOUS CARS AND THE LAW. 4 FEB 2015

GOODALL N. J. [2014]: Ethical Decision Making During Automated Vehicle Crashes, Transportation Research Record: Journal of the Transportation Research Board, pp. 58-65, 2014.

HOTTENTOT C.; MEINES V.; PINCKAERS M. [2015]: Experiments on autonomous and automated driving: an overview 2015. ANWB. The Hague. April 2015

WEINER G.; SMITH B. W. [2015]: Automated Driving: Legislative and Regulatory Action. Available at: http://www.cyberlaw.stanford.edu/wiki/index.php/Automated_Driving:_Legislative_and_Regulatory_Action

AUTONOMOUS | SELF-DRIVING VEHICLES LEGISLATION [2015]: available at: http://www.ncsl.org/research/transportation/autonomous-vehicles-legislation.aspx

BILL CA SB 1298: available at:

https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201120120SB1298

BILL DC B19-0931: available at:

http://dcclims1.dccouncil.us/lims/legislation.aspx?LegNo=B19-0931

BILL FL HB 1207: available at: https://www.flsenate.gov/Session/Bill/2012/1207

BILL FL HB 599: available at: https://www.flsenate.gov/Session/Bill/2012/0599

BILL FL SB 52: available at: https://www.flsenate.gov/Session/Bill/2013/0052

BILL MI SB 0169: available at:

http://www.legislature.mi.gov/%28S%28h2gyj51e4ahtwx21ikmjdi0z%29%29/mileg.aspx?page=getobject&objectname=2013-SB-0169

BILL MI SB 0663: available at:

http://www.legislature.mi.gov/%28S%28jqj5cob0o5n0et413jyuwnvy%29%29/mileg.aspx?page=GetObject&objectname=2013-SB-0663

BILL NV AB 511: available at:

https://www.leg.state.nv.us/Session/76th2011/reports/history.cfm?ID=1011

BILL NV SB 140: available at: http://leg.state.nv.us/76th2011/Reports/history.cfm?ID=324

BILL NV SB 313: available at:

http://leg.state.nv.us/session/77th2013/reports/history.cfm?ID=759

BILL ND HB 1065: available at: https://legiscan.com/ND/bill/1065/2015

BILL TN SB 598: available at: http://openstates.org/tn/bills/109/SB598/

List of Abbreviations

HAV	Highly Automated Vehicle
AD	Automated Driving
DoC	District of Columbia
DMV	Department of Motor Vehicles
DoT	Department of Transportation
EU	European Union
ANWB	Royal Dutch Touring Club
UK	United Kingdom
UNECE	United Nations Economic Commission for Europe
ITC	Intelligent Transport Systems
SAVI	Singapore Autonomous Vehicle Initiative
A*STAR	Agency for Science, Technology and Research
LTA	Land Transport Authority of Singapore
UVEK	Federal Department of Environment, Transport, Energy and Communications in Switzerland

List of Figures

List of Tables

Declaration concerning the Bachelor's Thesis

I hereby confirm that the presented thesis work has been done independently and using only the sources and resources as are listed. This thesis has not previously been submitted elsewhere for purposes of assessment.

Munich, September 24th, 2015

Diego Pacho Toubes

Technische Universität München - Lehrstuhl für Verkehrstechnik
Univ.-Prof. Dr.-Ing. Fritz Busch
Arcisstraße 21 80333 München, www.vt.bgu.tum.de

BACHELOR'S THESIS

State of the Art of Extraction of Traffic Messages from Social Media

Author:

Diego Pacho Toubes

Supervision:

Foteini Orfanou, M.Sc. (TUM)

Dr.-Ing. Miquel Estrada (UPC)

Date of Submission: 2015-11-06

Technische Universität München - Lehrstuhl für Verkehrstechnik
Univ.-Prof. Dr.-Ing. Fritz Busch
Arcisstraße 21 80333 München, www.vt.bgu.tum.de

BACHELOR'S THESIS

of Diego Pacho Toubes

Date of Issue: 2015-06-08

Date of Submission: 2015-11-06

Topic: **State of the Art of Extraction of Traffic Messages from Social Media**

Social media (Twitter, Facebook etc.) are widely used and connect people all around the world. More and more users of all age categories are joining the so called micro-blogging services for several purposes like communication and information or idea exchange. The use of social media has been expanded also towards the field of transportation. The users are sharing real time traffic related updates and messages such as incidents or traffic jam occurrences, special events that can affect traffic flow evolution and travel time, construction works, weather conditions etc.

Extraction of this kind of traffic information allows its reproduction and broadcast to a larger audience making the information much more useful. The fact that the traffic reports are in real time makes the extraction process more important and urgent as it will help road users planning their routes and avoid traffic problems.

The scope of this Bachelor Thesis is to analyze the current state of the art of the extraction of traffic messages from the various social media. More specifically the thesis will focus on the following aspects:

- Presentation of various traffic messages shared in social media.
- Presentation of the data which is extracted within the different approaches.
- Literature review and analysis of the existing tools, methods and techniques used for extracting and verifying traffic information from these sources.
- Literature review and analysis of the aims and purposes the extraction of such information serves (traffic management, traffic prediction etc.) and the validity of this information.

The student will present intermediate results to the supervisor (Foteini Orfanou, M.Sc.) in the fifth, tenth and 15th week.

The student must hold a 20 minute presentation with a subsequent discussion at the most two months after the submission of the thesis. The presentation will be considered in the final grade in cases where the thesis itself cannot be clearly evaluated.

____

Univ.-Prof. Dr.-Ing. Fritz Busch

Abstract

Micro-blogging is an emerging form of communication and has become very popular in recent years. Besides, micro-blogging services allow users to publish instant updates in short text messages that are broadcast to their followers or rest users in real-time. For this reason, Social Media users can post first-hand information when they are in traffic, but not driving, and if there is an anomaly, give information about its causes and effects. For this reason, every user could be seen as a mobile sensor, which, if correctly analyzed, would provide really useful information. But information extraction from these kinds of messages is a more challenging problem than from traditional trusted sources.

With this in mind, this paper presents a review of the main methods and techniques used to extract traffic information from Social Media. In addition, this paper is focused on an extensive literature to extract the main aims, purposes and validity of the different proposed approaches and techniques. The use of such approaches can develop a new vision of obtaining traffic related data from these sources, positively affecting the society as a whole. In addition, a SWOT analysis is carried out in order to test in which scenarios might be used such procedures to extract traffic information.

Table of Contents

1 Introduction

Social network sites, also called micro-blogging services (e.g., Twitter, Facebook, Tumblr), have spread in recent years, becoming a new kind of real-time information channel. Their popularity stems from the characteristics of portability thanks to several social networks applications for smartphones and tablets, easiness of use, and real-time nature. People intensely use social networks to report (personal or public) real-life events happening around them or simply to express their opinion on a given topic, through a public message. Such messages shared in social networks are also called Status Update Message (SUM) by D‘ANDREA ET AL. [2013],

On the other hand, social networks and media platforms have been widely used as a new source of information for the detection of events, such as traffic congestion, incidents, natural disasters (earthquakes, storms, fires, etc.) or other events. In particular, regarding traffic related events, people often share by SUMs information about the current traffic situation around them while driving. However, event detection from social networks analysis is a more challenging problem than from traditional media, where texts are well formatted. In fact, SUMs are unstructured and irregular texts. The main problems of such information shared in social media are: texts are typically very shorts due to characters' limitation, they usually contain noise, are written in an informal setting, containing, in many cases, incorrect sentences and the information in not reliable depending on the source.

For all of these reasons, and in order to analyze the information present in social networks, the use of text mining techniques, which employ methods from the fields of data mining, machine learning, statistics, and Natural Language Processing (NLP) to extract meaningful information has been thought as a new way to obtain valuable information. In other words, if such information is correctly analyzed, it may provide useful information about a traffic related event. In fact, social network users may be regarded as social sensors, open up new ways of obtaining traffic information in real time manner.

2 Related Work

With reference to current approaches for using social media to extract useful information for event detection, we need to distinguish between small-scale events and large-scale events. Small-scale events (e.g., traffic, car crashes, fires, or local manifestations) usually have a small number of SUMs related to them, belong to a precise geographic location, and are concentrated in a small time interval. On the other hand, large scale events (e.g., earthquakes, tornados, the election of a president) are characterized by a huge number of SUMs, and by a wider temporal and geographic coverage. Consequently and due to the smaller number of SUMs related to small-scale events, such event detection is a non-trivial task. Several works in the literature are related with event detection from social networks. However, while many works of the literature deal with large-scale event detection, only a few of them are focused on small-scale events.

Regarding large-scale event detection, SAKAKI ET AL. [2013] use Twitter streams to detect earthquakes and typhoons, by monitoring special trigger-keywords, and by applying an SVM as a binary classifier of positive events (earthquakes and typhoons) and negative events (non-events or other events). KRSTAJIC ET AL. [2012] present a method for detecting real-world events, such as natural disasters, by analyzing Twitter streams and by employing both NLP and term-frequency-based techniques. CHEW ET AL. [2010] analyze the content of tweets shared during the H1N1 (i.e., swine flu) outbreak, containing keywords and hashtags related to the H1N1 event to determine the kind of information exchanged by social media users. De LONGUEVILLE ET AL. [2009] analyze geo-tagged tweets to detect forest fire events and outline the affected area.

On the other hand, regarding small-scale event detection AGARWAL ET AL. [2012] focus on the detection of fires in a factory from Twitter stream analysis, by using standard neuro-linguistic programming techniques. F. ABEL ET AL. [2012], extract information from Twitter streams and merged it with information from emergency networks to detect and analyze small-scale incidents, such as fires. LI ET AL. [2012] propose a system, called TEDAS, to retrieve incident-related tweets. The system focuses on Crime and Disaster-related Events (CDE) such as shootings, thunderstorms, and car accidents, and aims to classify tweets as CDE events by exploiting a filtering based on keywords, spatial and temporal information, number of followers of the user, number of retweets, hashtags, links, and mentions.

3 Traffic Information in Social Media

Social media is great resource of user-generated contents. Public attention, opinion and hot topics can be captured in the Social Media, which provides the ability to predict human related events. Since Social Media can be retrieved in real time with no building and maintenance cost, such source could be probably used as a new sensor for those traffic management agencies, as a complement of their already available sensors on road. Social media has become an indicator of modern people and lifestyle in the Internet virtual community, as well as an indicator of people behaviour. Vast of user-generated contents strengthen linkage and interaction between each individuals within the online community, and also provide large amount of information related to various area.

In addition, and thanks to the booming of smartphones and cellular networks in the last years, it is easier to use and reach social media than ever before. Every day, more and more users are able to post and view contents on the net just simply using their handset devices. The trend of easy accessing social media will continuously grow with the development and commercialization of new wearable computer devices. On the other hand, traffic is one of the interesting and long-lasting problems in the world. For centuries, people worked to invent better tools and vehicles to easier travel and move goods. For recent decades, the intrigued traffic network has been established in almost all the big cities around the world, being responsible for both day-today traffic congestion and traffic anomalies, which significantly affects quality of life and impacts the economy. That is why understanding the dynamics of traffic has become an area of increasing interest.

Motivated by the potential value of Social Media and the difficulty of traffic flow prediction, there are several works in which their authors intend to build relationship between traffic information and social media data. Specially, the question of if social media will help to predict prior-event traffic is going to be answered. Furthermore, the purpose of research is to demonstrate and uncover both the great potential and value of big data, constantly generated by the social media. In Fig.3 we can find several examples of traffic related SUMs published in different Social Media (e.g., Twitter, Tumblr and Swarm).

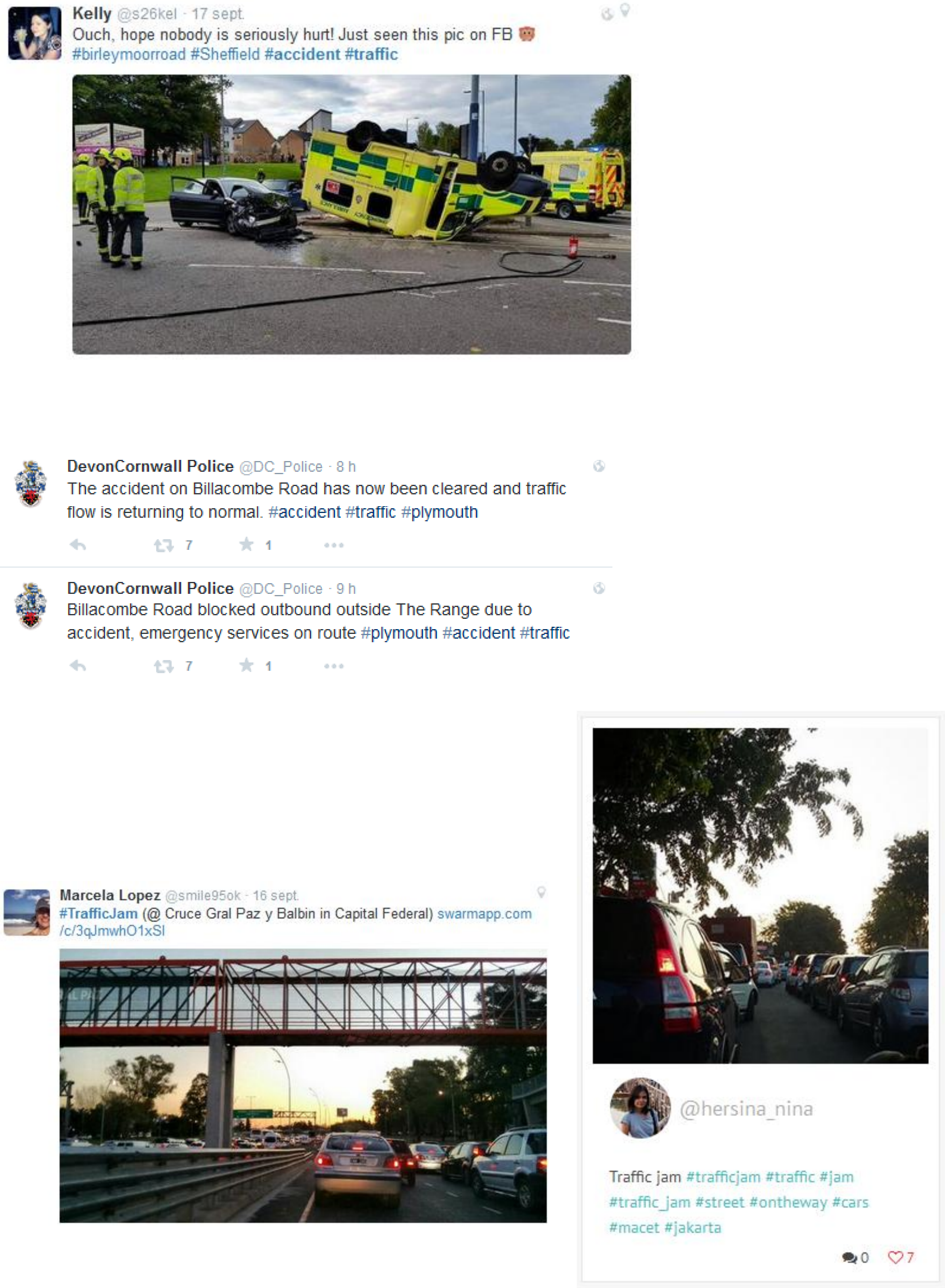

Fig.3.1 Traffic related SUMs in different Social Media

Source: Diego Pacho Toubes

4 Information Extraction Techniques

4.1 Literature Review

The use of messages posted in Social to obtain traffic related data is a new emerging research area, which –in some cases– has been implemented to traditional methods as new sources to obtain such information. Despite this we can find several works related to traffic data extraction performed with different approaches. WANICHAYAPONG ET AL. [2011] extract information from Twitter using syntactic analysis and then classifying it in two categories: point and link. Point is related only to one point on a road while link information associates with a road at its start and end point. The study was held in Thailand where more than a Million of tweets were collected for training and testing. The tweets are then tokenized using Lexto (dictionary based tokenizer with lexical analysis for Thai language) and the information extracted NLP techniques and syntactic analysis.

SAKAKI ET AL. [2012] extract, based on keywords, real-time driving information by analyzing Twitter's SUMs, and use a Support Vector Machine (SVM) classifier to filter "noisy" tweets not related to road traffic events. Schulz et al. [11] detect small-scale car incidents from Twitter stream analysis, by employing semantic web technologies, along with NLP and machine learning techniques. They perform the experiments using SVM, Naïve Bayes (NB), and RIPPER classifiers.

D'ANDREA ET AL. [2015] are focused on traffic detection and their aim is to detect and analyze traffic events by processing users' SUMs belonging to a certain area and written in the Italian language. They propose a system able to fetch, elaborate, and classify SUMs according to if they are related to a road traffic event or not. The traffic detection system was also employed for real-time monitoring of several areas of the Italian road network. They employ the SVM as a classification model by solving a binary classification problem (traffic versus non-traffic tweets). They also used C4.5 decision tree algorithm, NB and k-Nearest Neighbors (k-NN). They were also able to discriminate if traffic is caused by an external event or not, by solving a multiclass classification problem.

HE ET AL. [2011] examine whether it is possible to use social media information to improve longer-term traffic prediction. They first analyze the correlation between traffic volume and tweet counts with various granularities. Then they propose an optimization framework to extract traffic indicators based on tweet semantics using a transformation matrix, and incorporate them into traffic prediction, via linear regression. Experimental results using traffic and Twitter data originated from the San Francisco Bay area of California, demonstrate the effectiveness of their proposed framework.

MEGALLY [2012] presents a system that extracts traffic information from Twitter and designed to be used in route planning. She uses previously trained NB, Maximum Entropy (Maxent) and SVM classifiers to filter non relevant traffic. Then she applies Named Entity Recognition (NER) –Stanford NER– on traffic tweets to extract locations, highways and directions. These extracted locations are then geocoded and used in route planning to avoid routes with traffic jams.

ENDARNOTO ET AL. [2011] develope a system that can extract information of traffic from the Twitter account of TMC Polda Metro Jaya (police unit in Jakarta). Then they present such information in a map view by using Google Map and implementing it in Android-based mobile application. They use Part of Speech (PoS) and rules for information extraction constructed manually by analyzing random tweets of TMC.

ACHING ET AL. [2014] propose a strategy to use messages posted in a blogging platform for real-time sensing of traffic related information. Specifically, they use the data that appear in a blog, which is in Portuguese language and is managed by a Brazilian daily newspaper on its online edition. They propose a framework based on two modules to infer the location and traffic condition from unstructured, non georeferenced short posts in Portuguese. The first module relates to NER. It automatically recognizes three classes of named-entities from the input post (location, status and date) and then they apply a bootstrapping approach to expand the initially given list of locations, identifying new locations as well as locations corresponding to spelling variants and typographical errors of known locations.

NI [2013] uses social media information to assist traffic flow prediction under special event conditions. Specially, a short-term traffic flow prediction model, incorporated with tweet features, is developed to forecast incoming traffic flow prior to sport game events. She examines and compares the performance of four regression methods, autoregressive model, neural networks model, SVM, and k-NN respectively, with and without social media features.

MEILIN ET AL. [2011] present an application for Traffic Events Detection and Summary (TEDS) based on mining representative terms from those tweets posted when anomalies occur. Spatio-temporal analysis and a novel wavelet analysis model are applied for traffic event detection. They also use NLP to generate a cross correlation computation between tweets and traffic flow. Using the proposed signal processing technology, they demonstrate their system's effectiveness by examining traffic and metro travel in the area of Washington D.C.. Users of their application can search transportation status and analyze traffic events in specific locations.

ELSAFOURY [2013] proposes a system that is able to use traffic information shared by tweets in a real time manner. She uses a customized PoS tagging method to extract information from the tweets. Besides, PoS is used for geo-locating the tweets with custom developed locations' dictionaries. The result of the system is a map showing a highlighted route. This route is the location (road) mentioned in the tweet. The highlight or tone of the color depends on the traffic status which is also mentioned in the tweet.

CARVALHO ET AL. use micro-blogging messages posted on Twitter to perform real-time sensing of traffic-related information. They propose a text classification approach to automatically identify traffic-related messages posted on Twitter. They also want to create a suitable training set for setting up the classifier and drive it to a reasonable level of precision when identifying relevant messages. That is why they opt for a SVM to identify traffic related tweets and a dual-stage bootstrapping strategy.

TOSTES ET AL. analyze two social sensing sources, one derived from Foursquare, and another derived from Instagram, to verify if check-ins can be used as a hint of traffic conditions changes or current situation. Every data shared in Foursquare or Instagram is called check-in in their work. They want to investigate if they can use data obtained from social sensors (check-ins) to better understand the traffic condition. They use different algorithms to establish the correlation between check-ins and traffic flow in Manhattan, New York.

Finally, MCHUGH [2015] provides an approach of using big data extracted from Twitter, visualization and data mining techniques to predict and analyze traffic. This prediction model was used as an estimator to identify unusual traffic patterns in Dublin City. The generic model is designed using data mining techniques, multivariate regression algorithms, ARIMA and visually correlated with real-time traffic tweets.

4.2 System Architecture

How then does an information extraction system perform the kinds of complex processes required to identify and extract information? In general terms, an information extraction system is composed of a series of modules (or components) that process text by applying rules. Since information extraction involves selected pieces of data, an extraction system processes a text by creating computer data structures for relevant sections of a text while, at the same time, deleting irrelevant sections from the processing. Although there are variations among systems, generally the functions for the following set of modules are also performed somewhere in the processing.

4.3 Data Obtaining and Processing

The first step is the data obtaining and for this reason we need a device that provides us access to a large volume of information. Twitter allows access to a huge corpus of data through different Application Programming Interfaces, also called APIs (e.g., REST API, Search API and Streaming API). The Twitter Streaming API is used to retrieve tweets from public timelines, where these tweets are publicly available in real-time. We have to remember that through Twitter we can have access to data from other Social Media sites (e.g., blogs, Instagram, Foursquare) without the access restriction of their original sites. During this first module, we extract raw SUMs from the Twitter stream, based on one or more search criteria (e.g., geographic coordinates, keywords appearing in the text of the tweet). Each fetched raw SUM contains as follows: the user ID, timestamp, geographic coordinates, a retweet flag and the text of the tweet. The text may contain additional information such as hashtags, links, mentions and special characters.

After the SUMs have been fetched according to a specific search criterion, they are already pre-processed. In order to extract only the text of each raw tweet, removing then all meta-information associated with it, a Regular Expression filter should be applied. More in detail, the most common meta-information discarded are: user id, timestamp, geographic coordinates, hashtags, links, mentions, and special characters. Finally, a case-folding operation is applied to the texts, in order to convert all characters to lower case. At the end of this elaboration, each fetched SUM appears as a string, that is to say, a sequence of characters.

The following processing module is performed to transform the set of pre-processed SUMs in another of numeric vectors to be elaborated by the Classification of SUMs module. To this aim, some text mining techniques are applied in sequence to the pre-processed SUMs. In the following, the text mining steps performed in this module are described in detail. Besides, such steps are used by several authors of the literature, but these that are presented below are specifically described and used by D'ANDREA ET AL. [2015]:

a) *tokenization* is typically the first step of the text mining process, and consists in transforming a stream of characters into a stream of processing units called tokens (e.g., syllables, words, or phrases). During this step, other operations are usually performed, such as removal of punctuation and other non-text characters, and normalization of symbols (e.g., accents, apostrophes, hyphens, tabs and spaces). The tokenizer removes all punctuation marks and splits each SUM into tokens corresponding to words (also called bag-of-words representation). At the end of this step, each $\mathrm{SUM_j}$ is represented as the sequence of words contained in it. The jth tokenized SUM is described as $SUM_j^T = \left\{t_{j1}^T, \dots, t_{jh}^T, \dots, t_{jH_j}^T\right\}$, where t_{jHj}^T is the hth token and H_j is the total number of tokens in SUM_j^T;

b) *stop-word filtering* consists in eliminating stop-words or those words which provide little or no information to the text analysis. Common stop-words are articles, conjunctions, prepositions, pronouns, etc. Other stop-words are those having no statistical significance, that is, those that typically appear very often in sentences of the considered language (language-specific stop-words), or in the set of texts being analyzed (domain-specific stop-words), and therefore can be considered as noise. Several authors have shown that the 10 most frequent words in texts and documents of the English language are about the 20–30% of the tokens in a given document. At the end of this step, each SUM is thus reduced to a sequence of relevant tokens. The jth stop-word filtered SUM is described as $SUM_j^{SW} = \left\{t_{j1}^{SW}, \ldots, t_{jk}^{SW}, \ldots, t_{jK_j}^{SW}\right\}$, where t_{jk}^{SW} is the kth relevant token and K_j , with $K_j \leq H_j$, is the total number of relevant tokens in SUM_j^{SW}. As seen before, a relevant token is that which does not belong to the set of stopwords;

c) *stemming* is the process of reducing each word or token to its stem or root form by removing its suffix. The purpose of this step is to group words with the same theme having closely related semantics. Hence, at the end of this step, each SUM is represented as a sequence of stems extracted from the tokens contained in it. The jth stemmed SUM is described as $SUM_j^S = \left\{t_{j1}^S, \ldots, t_{jl}^S, \ldots, t_{jL_j}^S\right\}$, where t_{jl}^S is the lth stem and L_j , with $L_j \leq K_j$, is the total number of stems in SUM_j^S ;

d) *stem filtering* consists in reducing the number of stems of each SUM. In particular, each SUM is filtered by removing from the set of stems the ones not belonging to the set of relevant stems. The set of $F\ relevant\ stems\ RS = \left\{\hat{s}_1, \ldots, \hat{s}_f, \ldots, \hat{s}_F\right\}$ is identified during the learning stage. At the end of this step, each SUM is represented as a sequence of relevant stems. The jth filtered SUM is described as $SUM_{Fj}^S = \left\{t_{j1}^{SF}, \ldots, t_{jp}^{SF}, \ldots, t_{jP_j}^{SF}\right\}$, where $t_{jp}^{SF} \in RS$ is the pth relevant stem and P_j, with $P_j \leq L_j$ and $P_j \leq F$, is the total number of relevant stems in SUM_{Fj}^S ;

e) *feature representation* consists in building, for each SUM, its corresponding vector of numeric features. Indeed, in order to classify the SUMs, we have to represent them in the same feature space. In particular, an F-dimensional set of features $X = \left\{X_1, \ldots, X_f, \ldots, X_F\right\}$ are considered, corresponding to the set of relevant stems. For each SUM_{Fj}^S a vector $x_j = \left\{x_{j1}, \ldots, x_{jf}, \ldots, x_{jF}\right\}$ is defined, where each element is set according to the following formula:

$$x_{jf} = \begin{cases} w_f\ if\ stem\ \hat{s}_f \in SUM_{Fj}^S \\ 0\ otherwise \end{cases} \tag{4}$$

In (4), w_f is the numeric weight associated to the relevant stem $\hat{s}_f$.

In Fig.4.1 we can see, step by step, all the processes applied to a sample tweet as described above.

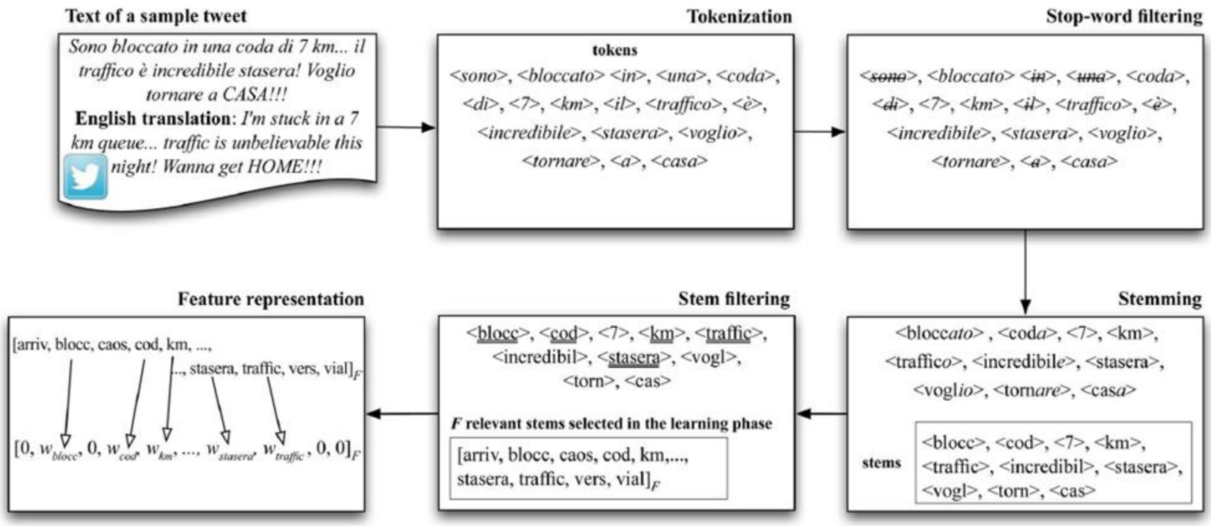

Fig.4.1 Steps applied to a sample tweet

Source: D'ANDREA ET AL. [2015]

4.4 Information Classification

The next module is based on assigning to each elaborated SUM a class label related to traffic events. Applying information extraction on text is linked to the problem of text simplification (classification) in order to create a structured view of the information present in free text. That is why, after the information extraction module, we need to classify all the SUMs depending on their content. This process is the most important and complex one because it must determine whether a SUM is considered important (traffic related) or noise (not traffic related). There are two main information extraction methods used for such purpose: Named Entity Recognition (NER) and Part of Speech (PoS).

4.4.1 Named Entity Recognition

Usually the recognition task involves assigning a unique identifier to each extracted entity. In addition, entities are discrete objects and the task of NER system is to find and classify these entities to then, enable Information Extraction without having any existing knowledge about the entity meaning.

Then, the first step is based on training NER system using classifiers to recognize those entities that we want to extract from SUMs. The two main classes of entities that have been continually found in the literature, and then are considered as relevant are described as follows:

- *location* is related to words that indicate the names of streets, roads, highways, bridges, etc;

- *status* is related to words that indicate the traffic status on a location (e.g., blocked, congested, closed).

The second phase is supported by an especially built dictionary that contains frequently occurring words in SUMs related to traffic. Then we just have to extract this information classified according to each entity.

4.4.1.1 Sytem Training

In order to train our NER system to extract traffic information from SUMs, we need to use classification models. The task of text classification –as mentioned above– is to determine, for an object, to which class of a given set it belongs. It is then when a decision should be taken concerning each incoming SUM. All of them have to be classified into one of two classes depending on: *traffic* or *non-traffic*. A *traffic* SUM is one that mentions any information related to traffic while *non-traffic* refers to any irrelevant tweet. This process can be done manually or using a machine learning approach, distinguishing between whether it is either supervised or unsupervised.

Supervised learning is the machine learning task of inferring a function from labeled training data. We need a set of good example documents referred as training documents. The training dataset consists of a collection of labeled documents, where labeling refers to the process of annotating each document with its class. Using a learning method we wish to learn a classifier to map documents to classes. This is also called supervised learning because the process is supervised by a human who assigns the labels. In supervised learning each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). An optimal scenario will allow the algorithm to correctly determine the class labels for unseen instances.

Now, classifiers can be classified according to two different categories: generative or discriminative. Generative models are those that generate the observed data from hidden states. In a generative model, the focus is on the joint probability and our goal is to maximize such joint likelihood. In contrast, discriminative models target the classification decision that we want to make, taking the data as given and putting a probability over a hidden structure. Discriminative models have been widely used due to they tend to have high accuracy performance and they are linguistically interesting because of the ease of adding lots of linguistic features. The discriminative model is focused on the conditional probability and seeks to maximize it. Therefore, the fundamental difference between these two models is that discriminative learn the boundary between classes while generative model the distribution of individual classes. Now the main supervised classifiers used in the literature are introduced as follows.

- [1]**Naïve Bayes** (NB) is a supervised probabilistic learning method. It is based on the bag-of-words model. The bag-of-words representation neglects every information about the position of a word in a document, and instead, focuses only on the word frequency in the document. Each document will be represented as a set of words and their frequency count.

The conditional probability $P(c|d)$ for a document *d* and a class *c* according to Bayes rule is:

$$P(c|d) = P(c) \cdot \frac{P(d|c)}{P(d)} \tag{4.1}$$

Where $P(c)$ is the probability of the class *c*, $P(d|c)$ is the probability of document *d* given the class *c* and $P(d)$ is the probability of document *d*. Each document *d* is presented as a set of words $w_1, w_2, w_2, \ldots, w_k$. We estimate the probability $P(d|c)$ of a document *d*, given a class *c*, as the multiplication of the probabilities of each individual word w_i given the class *c*.

$P(d|c)$ can then be written as follows:

$$\begin{aligned} P(d|c) &= P(w_1, w_2, w_3, \ldots, w_k|c) \\ &= P(w_1|c) \cdot P(w_2|c) \cdot P(w_3|c) \cdot \ldots \cdot P(w_k|c) \\ &= \prod_{w_i=1,2,\ldots,k} P(w_i|c) \end{aligned} \tag{4.2}$$

However, the probability $P(w_i|c)$ might be equal to zero. To eliminate zeros, we use add-one or Laplace smoothing, which simply adds one to each word w_i count. Add-one smoothing can be interpreted as a uniform prior (each term occurs once for each class) that is then updated as evidence from the training data coming in.

Then the probability of a class *c* needs to be computed. This is equal to the relative frequency of the class *c* in the corpus of training documents. It denotes how often this class occurs.

$$P(c) = \frac{N_C}{N} \tag{4.3}$$

Where N_C is the number of documents in class *c* and *N* is the total number of documents.

[1] *M. MEGALLY [2012]: Information Extraction from Social Media for Route Planning. Master's thesis Nr. 3413, Universität Stuttgart, November 15, 2012.*

- The [1]**Maximum Entropy** (Maxent) classifier is a discriminative classifier. During the Maxent classification, each document *d* is represented as a set of features $f_1, f_2, \ldots, f_k$. The model calculates a linear function of these features. This linear function represents a score for each class and can be written as:

$$vote(c) = \sum_i \lambda_i \cdot f_i(c,d) \tag{4.4}$$

Where λ_i is a weight assigned to each feature f_i . $vote(c)$ is calculated for each class c and the class that maximizes it is then chosen.

- [1]**Support Vector Machine** (SVM) is another example of a discriminative model. It is a vector space based machine learning method, where its objective is to find a decision boundary between two classes, which is maximally far from any point in the training data. There might exist several possible linear separators between the two classes in the training data set. However, it is always better to choose a decision boundary that is as far as possible from all data points of the two classes. SVM adopts this approach in choosing a decision boundary in binary classification. It defines a criterion to pick a decision surface that is maximally far away of any data point. This distance is referred to as the *margin* of the classifier. The construction of the margin mainly relies on the selection of data points known as support vectors that specifically determine the position of the separator. The maximized margin is shown in Figure 4.2. The large margin ensures the decrease of low certainty classification decisions. This is referred to as the *classification safety margin*.

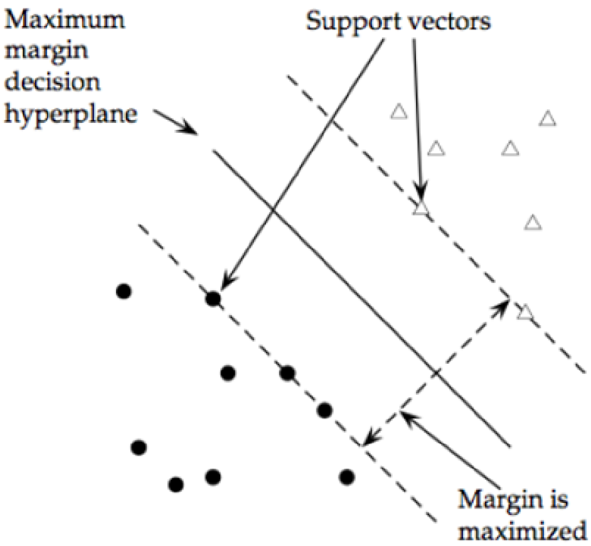

Fig.4.2 Separating hyperplane in SVM.

Source: MEGALLY [2012]

During the SVM construction, an SVM classifier insists on the choice of a large margin around the decision boundary. A decision hyperplane can be defined by an intercept term *b* and a decision hyperplane normal vector $\vec{w}$ which is perpendicular to the hyperplane. This vector is called the weight vector. All points $\vec{x}$ on the hyperplane satisfy $\vec{w}^T\vec{x} = -b$. The linear classifier is then:

$$f(\vec{x}) = sign(\vec{w}^T\vec{x} + b) \quad (4.5)$$

The algorithm starts with a training data set, $D = \{(\vec{x_i}, y_i)\}$, where each point $\vec{x_i}$ has a class label y_i. For a binary SVM the class labels are always +1 and -1. The data defines the best separating hyperplane and the data is fed into a quadratic optimization procedure to find this plane.

- [2]**Decision tree learning** (DT) is also a discriminative method, commonly used in data mining. It main objective, again, is to create a model that predicts the value of a target variable based on several input variables. A decision tree is a simple representation for classifying examples. In addition, decision tree learning is one of the most successful techniques for supervised classification learning.

For this section, the assumption that all of the features have finite discrete domains is considered, and there is a single target feature called the classification. Each element of the domain of the classification is called a class. A decision tree or a classification tree is a tree in which each internal (non-leaf) node is labeled with an input feature. The arcs coming from a node labeled with a feature are labeled with each of the possible values of the feature. Each leaf of the tree is labeled with a class or a probability distribution over the classes. A tree can be "learned" by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node has all the same value of the target variable, or when splitting no longer adds value to the predictions. This process of top-down induction of decision trees (TDIDT) is an example of a greedy algorithm, and it is –by far– the most common strategy for learning decision trees from data.

- Finally, the [2]**k-Nearest Neighbors algorithm** (k-NN) is also a discriminative algorithm, since it models the conditional probability of a sample belonging to a given class. k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The *k*-

[2] *Jeffrey Strickland [2015]: Predictive Analytics using R. Lulu.com*

NN algorithm is among the simplest of all machine learning algorithms. It can be useful to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the classification phase, *k* is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label that is most frequent among the *k* training samples nearest to that query point. A commonly used distance metric for continuous variables is Euclidean distance, for example for discrete variables, such as for text classification. Often, the classification accuracy of k-NN can be improved significantly whether the distance metric is learned with specialized algorithms.

However, we also find unsupervised classifiers that automatically identify traffic-related messages without any human supervision. These methods are not as popular as the supervised, and very few forks are performed using unsupervised classifiers. Moreover, they are also combined with supervised methods and are also used after the supervised system is enough reliable.

4.4.2 Part of Speech

A [3]Part of Speech is a category of words, or more generally, of lexical items, which have similar grammatical properties. Words that are assigned to the same part of speech generally display similar behavior in terms of syntax and sometimes in terms of morphology. Commonly listed English parts of speech are noun, verb, adjective, adverb, pronoun, preposition, conjunction, interjection, and sometimes numeral, article or determiner.

A part of speech may also be called a word class, lexical class, or lexical category, although the term lexical category refers in some contexts to a particular type of syntactic category, and may thus exclude parts of speech that are considered to be functional, such as pronouns. The term form class is also used, although this has various conflicting definitions. Word classes may be classified as open or closed: open classes (like nouns, verbs and adjectives) acquire new members constantly, while closed classes (such as pronouns and conjunctions) rarely acquire new members.

Almost all languages have the word classes, noun and verb, but beyond these there are significant variations in different languages. For example, Japanese has as many as three

[3] *https://en.wikipedia.org/wiki/Part_of_speech*

classes of adjectives where English has one; Chinese, Korean and Japanese have a class of nominal classifiers; many languages lack a distinction between adjectives and adverbs, or between adjectives and verbs. This variation in the number of categories and their identifying properties means that analysis needs to be done for each individual language. Nevertheless, the labels for each category are assigned on the basis of universal criteria.

According to Elsafoury [2013] Part of Speech tagging (PoS) is basically classifying each token that is produced by the previous step, based on pre-defined word class or PoS name. Using rules that were previously defined in the database, tagging process will search whether each of the token match with one of the word in the vocabulary collection in the database for each possible word class or PoS name.

PoS tagger is a computer program that does this task. Taggers use a large amount of annotated training corpus to tag (label) properly. There are three main approaches used for PoS taggers. These types are rule-based, stochastic or transformation-based learning approaches. The difference between these approaches is how they assign the tags to the words.

- *Rule-based approach* uses dictionary or lexicon to assign tags to words. Hand write rules are used to select the most suitable tag when there are more than one suggested tags. This approach needs a direct human interaction to check the written rules.

- The *stochastic approach* also called probabilistic, uses a training corpus to assign the most suitable tag for a word. Stochastic taggers use Hidden Markov Model (HMM). Markov model is a machine learning method based on probabilistic models. HMM is used to find the optimal tags sequence $T = \{t_1, t_2, \dots, t_n\}$ for the given words sequence $W = \{w_1, w_2, \dots, w_n\}$.

- The *transformed-based approach* is a mixture between the rule-based approach and the stochastic approach. It tags the given text automatically but based on rule based algorithm. The transformation-based approach picks up the most probable tag based on a training corpus then applies a set of written rules to see how suitable the chosen tag is.

4.5 Information Extraction

Now we have to adopt a template-based information extraction approach (IE). This approach requires predefined template schemas and labeled data to learn to extract their slot fillers. Each template has several slots and the information we extract is used to fill these slots. The template extraction algorithm requires full knowledge of the templates and labeled corpora, such as in rule-based systems. Furthermore, template design for information extraction depends on the nature of the task and therefore it affects the success of capturing information from text. The template has to be clear, deterministic, descriptive and monotonic.

Moreover, the information extraction depends on the task used to classify it, and that is why it is necessary to distinguish between NER and PoS. For each method we have to define different relationships between words in order to obtain the final output such as identifying an accident or a traffic jam.

4.5.1 Named Entity Recognition

After annotating traffic tweets and recognizing named entities that appear in them, an IE module is responsible for filling the fields of a template for every SUM. By the analysis of traffic SUM performed by D‘ANDREA ET AL. [2015], it was found that when a traffic jam or an accident is reported at a particular place, people usually mention the location by specifying a start and end point of the traffic jam. Social Media users usually also mention a specific landmark or road intersections that are blocked, and this is referred to as ref-point in the template. Furthermore, if the traffic information is published by a traffic channel, it usually includes the city or county where it happens. Additionally, when a highway is mentioned in a tweet, it is accompanied by a direction that is also extracted to fill a slot in the template. The fields of the template are generally as follows:

- *start point* of the traffic incident;

- *end point* of the traffic incident;

- *highway(s)*: one or more highways mentioned in the tweet where the incident occurs;

- *status*: is related to words that indicate the traffic status on a street.

These slots are filled using previously defined rules along with the annotations or token types attached to each token. The slots are optional, so not all the information have to be extracted from a single tweet, partly owing to SUMs structure and their limitation of characters in text.

The fist example present in the literature is performed by MEGALLY [2012]. She has implemented the following rules based on an analysis of 1000 tweets to fill the following described template slots:

- *start point*: any token labeled as location or highway and occurs after a ”from” or ”between” preposition is considered a starting point.

- *end point*: any token labeled as location or highway and occurs after ”to” or ”onto” or the ”between ... and” preposition is considered a starting point.

- *ref-point*: any token labeled as location or highway and occurs after these prepositions: after, past, near, at, before, beyond, across

- *city*: a token labeled as location that occurs as the first token in the tweet or after ”in” preposition.

- *highway and direction*: any token labeled as highway that occurs after ”on” preposition or is not preceded by any prepositions is used to fill this slot along with the direction if it’s mentioned.

An example of a tweet and its template slots extracted according to above described rules is:

“@SUNOL: Accident NB 680 past Koopman Rd has all lanes blocked as a medivac helicopter is landing on the freeway. Expect delays.”

The template filled out of this tweet is as follows:

- Ref-point : KOOPMAN ROAD

- City: SUNOL

- Highway: 680

- Direction: NB

ACHING ET AL. [2014] identified a set of patterns that are easily recognizable, which occur frequently and across several posts, and that indisputably indicate the relations that can be used to extract the location and status information contained in the post. These relations are:

- *restriction relation*: a binary relation that declares a proximity relation between two location entities;

- *reference relation*: a ternary relation that declares the status and direction of a specific location entity.

If we denote the location entities as LOC_1 , LOC_2, and so on, and status entities as $STATUS$, then, restriction and reference relations can be defined by the indicator functions:

$$Restriction\ (LOC_1\ , LOC_2) \tag{4.5.1.1}$$

$$Reference\ (LOC_1\ , LOC_2, STATUS) \tag{4.5.1.2}$$

As an example of use of the IE module, consider the following tagged post, shown in Fig.4.3.

$$\underbrace{Brasil\ Avenue}_{LOC_1}\ direction\ \underbrace{Ibirapuera\ Avenue}_{LOC_2}$$
$$with\ \underbrace{slow\ traffic}_{STATUS}\ from\ \underbrace{Venezuela\ Street}_{LOC_3}$$
$$until\ \underbrace{Pedro\ Alvares\ Cabral\ Avenue}_{LOC_4}$$

Fig.4.3 Sample tweet

Source: ACHING ET AL. [2014]

First, the algorithm identifies the ternary relation that satisfies the pattern: LOC_1 + direction + LOC_2 + with + $STATUS$. In this case, $Reference$ *(Brasil Avenue, Ibirapuera Avenue, slow traffic)* satisfies that relation. Second, the algorithm identifies the binary relation that satisfies the pattern: from + LOC_3 + until + LOC_4 as for instance, $Restriction$*(Venezuela Street, Pedro Alvares Cabral Avenue)*. This way, the output tuple is: $T = (t_1, t_2, t_3, t_4, t_5)$ = *(Brasil Avenue, Ibirapuera Avenue, slow traffic, Venezuela Street, Pedro Alvares Cabral Avenue),* where the tuple T contains the entities related to the information of location (t_1), direction (t_2), status (t_3), and physical restriction (t_4 and t_5), extracted from the input post.

Another proposed methodology is stated by WANICHAYAPONG ET AL. [2011], with the purpose to extract the name of the road and its start and end point to classify the extracted traffic information.

- *Road attribute extraction:* they compare each token with the *Road Word of Place* dictionary. If that token matches a road entry in database, that token will be marked as a road attribute for this traffic information. Then, they parse the rest of traffic information tokens during the next step: start point attribute extraction.

- *Start point attribute extraction*: they look amongst the tokens for a start point attribute by comparing each word with the *Start Preposition* and *Place* dictionary. If that word matches a start preposition in the dictionary and its following word matches a place in the dictionary, that word that matches a place will be marked as a start point attribute of this traffic information.

- *End point attribute extraction*: similarly to start point attribute extraction, they look amongst an end point attribute by comparing each word with the *End Preposition* and *Place* dictionary. If that word matches an end preposition in the dictionary and its following word matches a place in the dictionary, that word that matches a place will be marked as an end point attribute of this traffic information.

- *Classification*: *link* information is the traffic information that has filled all the attributes: road, start point and end point. *Point* information is the traffic information that has road and start or stop point attribute or only road, start point attribute or stop attribute, as shown in Tab.4.1.

Message	Road	Start	End	Category
สุขุมวิทขาออกข้ามแยกบางนามาหน้ากรมอุตุฯติด	สุขุมวิท	แยกบางนา	กรมอุตุฯ	Link
แจ้งวัฒนะฝนตกแล้วจร้า	แจ้งวัฒนะ	-	-	Point
ส.กรุงธนฯเข้าท้ายชะลอตัวในถนนสิรินทร	ส.กรุงธนฯ	-	ถนนสิรินทร	Point
ส.กรุงธนฯ มุ่งหน้าอนุสาวรีย์ฯ รถโล่ง	ส.กรุงธนฯ	อนุสาวรีย์ฯ	-	Point

Tab.4.1 Traffic news filtering sample

Source: WANICHAYAPONG ET AL. [2011]

4.5.2 Part of Speech

Once each token has been tagged, those tokens will be analyzed by the system using some pre-defined rules, which are basically gathered from the analysis of tweet format. Those rules will be used to gather useful information regarding traffic, such as time, origin, destination and traffic condition. Then, these 4 parameters will be stored in the database. Those attributes of information that we want to extract are basically template, and during this step the database will be filled with the information that is extracted from the tweet.

Fig.4.4 shows three different examples performed by ENDARNOTO ET AL. [2011] in which they analyze tweets from the Traffic Management Center (TMC) of Jakarta. The first example means that at 06:39, from Fatmawati (name of place) to Blok A (name of place) traffic crowded nearly jammed. The second means that at 06:23, from Joglo Main Road to Pos Pengumben (name of place) traffic jammed. And finally, the third example means that at 10:51, from Bekasi (name of place) to Cawang (name of place) traffic crowded nearly jammed. PoS Tag phase uses Tagset, as shown in Tab.4.2, and rules for information extraction that are constructed manually by analyzing random tweets of TMC.

```
06:39 Fatmawati arah ke Blok A padat merayap.
06:23 Jl.Raya Joglo arah Pos Pengumben lalin tersendat
10:51 Bekasi arah Cawang lalin padat merayap
```

Fig.4.4 Sample of TMC's tweet

Source: ENDARNOTO ET AL. [2011]

No	POS	POS Name	Example
1	AJ	Adjective	Ramai (crowded), Macet (jammed)
2	AT	Adjective Time	06:50
3	AV	Adverb	Sangat (highly)
4	CJ	Conjunction	Dan (and), Lalu (then)
5	N	Noun	Lalin (traffic), Arus (stream)
6	NP	Noun Place	Pondok Indah, Bintaro
7	P	Preposition	Di (at), Ke (to), Dari (from)
8	V	Verb	Merayap (crawling), Terjadi (happening)

Tab.4.2 Meaning of each tweet's word

Source: ENDARNOTO ET AL. [2011]

Fig.4.5 shows three examples of rules that they used in their experiment. Those rules represent the sequence of appearance of PoS name in a tweet. For example, if we know that the sequence matches the first rule (AT NP P NP AJ V) or third rule (AT NP P NP N AJ V), then we can get AT (adjective time) as time, first NP (noun place) as origin, second NP as destination and AJ+V (adjective followed by verb) as traffic condition. Also for the second rule, we can get AT as time, first NP as origin, second NP as destination and AJ as traffic condition.

```
AT NP P NP AJ V
AT NP P NP N AJ
AT NP P NP N AJ V
```

Fig.4.5 Example of rules

Source: ENDARNOTO ET AL. [2011]

```
Time : 06:39
From : Fatmawati
To : Blok A
Condition : padat merayap
```

Fig.4.6 Result from the system

Source: ENDARNOTO ET AL. [2011]

For her part, ELSAFOURY [2013] shows another example of how to extract traffic information applying PoS tagging. The role of PoS tagging is to generate the tagged set of the input tweets. Then, and in order to extract information the tagged set are compared against a pre-defined dictionaries for traffic status and cause, as shown in Fig.4.7.

To extract traffic status from tweets, the proposed system extracts the tagged set from the input tweets. Then the tagged set is filtered to get all the nouns, verbs, adjectives and adverbs out of the tagged set. Then, these words are compared with traffic status dictionaries. Such dictionaries are three different depending on if traffic condition status is normal, crowded or jammed. The same methodology is used to extract the cause of such traffic condition. The proposed system extracts all nouns from the tagged set, and then these nouns are compared with the cause dictionary, in order to know the reason behind such traffic status.

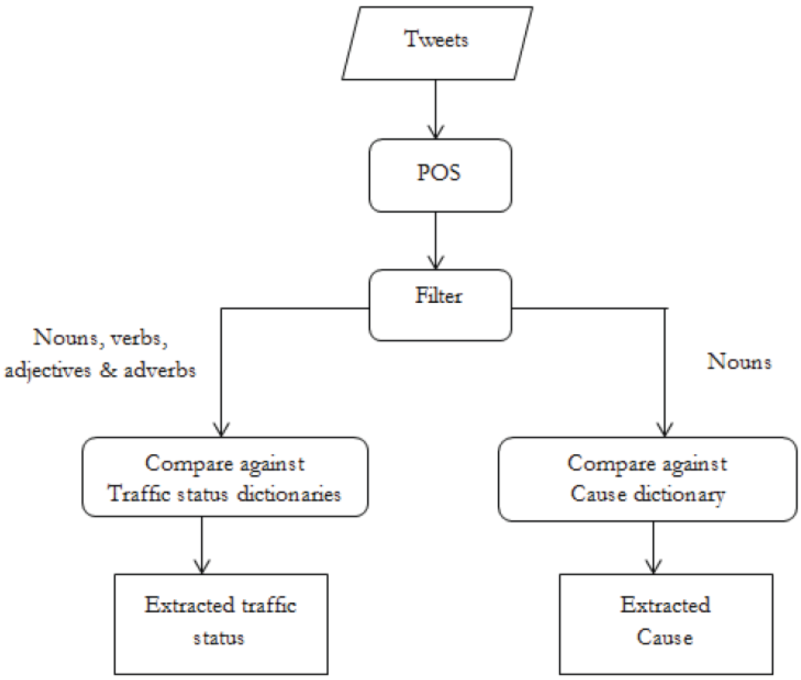

Fig.4.7 Information extraction

Source: ELSAFOURY [2013]

In her research, the system needs four main entities of information to be extracted from tweets. These entities are as follows:

- *location*: to extract the location entity (address), all the nouns were extracted from the tweets and, then these nouns were compared to a list of London's locations dictionary which was saved in a local text file. This list contains all the streets covered by "Central London" corridor. If there are different locations list in the processed tweets, then these locations are taken and separated by "/";

For example, the next tweet is talking about an incident in a particular location.

"The A4 Ellesmere Rd has reopened at Sutton Court Rd following the earlier collision. Residual Qs remain back to junction 2 on the M4"

the output for this step is <a4ellesmererd/ Sutton court rd>

- *road segments*: if the extracted streets names has a composite name, this location is compared against "Road_Segmnts" dictionary. If it exists then street name will be replaced in the output with its latitude and longitude;

the output of this step is <51.487570, -0.264304/ Sutton court rd>

- *estimation locations*: they are locations, e.g., cross roads intersection name, mentioned in the tweets to report more accurate incident location. These estimation locations or distances usually come after preposition words e.g. near, near to, close to, about, at, in front of, behind or from. Firstly, all the preposition words were extracted from the tweets. Then, the words were extracted after these prepositions. If there are more than one estimated location, they will be extracted and separated by "/". If the preposition "at" exists in the extracted prepositions tweet text, then the extracted locations from the first step will be compared against "Roads_Intersections" dictionary. If the locations exist in the dictionary, then the latitude and longitude of the intersection will be retrieved;

the output of the location's step is < 51.487596,-0.267255>

- *traffic status*: to extract the traffic status, all nouns, verbs, adjectives and adverbs were extracted from tweets. Then, it was compared to traffic keywords saved in three text files " Jammed_Traffic", "Crowded_Traffic" and "Normal_Traffic". The first file contains keywords related to jammed traffic e.g. closed, blocked, blocking or jam. The second one contains words related to crowded e.g. delays, crowd, crowded or slow and the last one keywords related to normal traffic e.g. normal, clear, smooth or open;

the output of this step is <normal>

- *cause*: to extract the causes from the tweets all nouns are extracted from the tweets. Then compare it to pre-defined causes keywords saved in a local text file “causes”. This file contains keywords like accident, crash, roadworks or flood;

the output of the cause is: <collision>

The full output of the algorithm for the pre-mentioned tweet is as follows:

- location <51.487570, -0.264304/ Sutton court rd>
- estimation location <51.487596,-0.267255>
- traffic status <normal>
- cause <collision>

5 Evaluation in Traffic Information

Given an input text, or a collection of texts, the expected output of an information extraction system can be defined very precisely, facilitating the evaluation of different information extraction systems and approaches. In order to evaluate the results achieved by such information extraction systems, for main statistical metrics are widely used. Such statistical are precision, accuracy, recall, and F-score. With the objective of explaining the meaning of every single metric, we are going to use the method of binary classification, which distinguish between positive and negative class. In fact, in the case of a multi-class classification, the metrics are computed per class and the overall statistical measure is simply the average of the per-class measures.

The correctness of a classification can be evaluated according to four values: 1) true positives (TP) to the number of real positive samples correctly classified as positive; 2) true negatives (TN) to the number of real negative samples correctly classified as negative; 3) false positives (FP) to the number of real negative samples incorrectly classified as positive; and finally 4) false negatives (FN) to the number of real positive samples incorrectly classified as negative.

Based on the previous definitions, we are now able to define the employed statistical metrics and provide their corresponding equations. Accuracy (5.1) represents the overall effectiveness of the classifier and corresponds to the number of correctly classified samples over the total number of samples. For its part, Precision (5.2) is the number of correctly classified samples of a class, in other words, positive class over the number of samples classified as belonging to that class. Recall (5.3) is the number of correctly classified samples of a class, that is to say, positive class, over the number of samples of that class; it represents the effectiveness of the classifier to identify positive samples. Thus, recall and precision can be seen as measure of completeness and correctness, respectively. Finally, the F-score (5.4) (typically used with $\beta = 1$ for class-balanced datasets) is the weighted harmonic mean of precision and recall, and it is generally used to compare different classifiers.

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{5.1}$$

$$Prec = \frac{TP}{TP + FP} \tag{5.2}$$

$$Rec = \frac{TP}{TP + FN} \tag{5.3}$$

$$F_{\beta-score} = (1 + \beta^2) \cdot \frac{Prec \cdot Rec}{\beta^2 \cdot Prec + Rec} \tag{5.4}$$

We found several examples of accuracy's values between the different works performed in the literature, as shown below. WANICHAYAPONG ET AL. [2011] obtained an accuracy of 91.75% by using an approach that considers the presence of place mentions and special keywords in the tweet. This method can classify traffic tweets into the point category with 76.85% accuracy and traffic tweets into the link category with 93.23% accuracy. LI ET AL. [2012] achieved an accuracy of 80% for detecting incident-related tweets using Twitter specific features, such as hashtags, mentions, URLs, and spatial and temporal information. SAKAKI ET AL. [2012] employed an SVM to identify heavy-traffic tweets and obtained an accuracy of 87%.

SCHULZ ET AL. [2013], by using SVM, RIPPER, and NB classifiers, obtained accuracies of 89.06%, 85.93%, and 86.25%, respectively. D'ANDREA ET AL. [2015] achieved an average accuracy of 95.75% using SVM distinguishing between traffic or non-traffic related tweets and an accuracy of 88.89% (being SVM again the best classifier) to discriminate the cause of traffic. VERMA ET AL. [2011] reported on an approach of using linguistic features to detect tweets with content relevant to situational awareness during mass emergencies (emergency event detection), which achieved 80% accuracy. BENSON ET AL. [2011] presented a CRF-based approach to extracting entertainment events with 85% precision through aggregating information across multiple messages. LOCKE ET AL. [2009] shown results of tuning a SVM-based classifier for classifying persons, locations and organizations names in Twitter, which achieves relatively poor results, as precision and recall around 75% and 50% respectively.

RITTER ET AL. [2011] described an approach to segmentation and classification of a wider range of names in tweets based on CRFs (using POS and shallow parsing features) and Labeled Latent Dirichlet Allocation respectively. The reported precision oscillated around 70%. MEGALLY [2012] combined the three classifiers (NB, Maxent and SVM) in a voting system achieving an accuracy of 99.4% for a first classification for each incoming tweet into either traffic or junk. Additionally, she classified each traffic tweet into positive or negative. SVMs outperformed NB and Maxent in this classification type achieving an accuracy of 94.06%. Finally, ACHING ET AL. [2014] used a first bootstrapping approach to expand the initially given list of locations, identifying new locations as well as locations corresponding to spelling variants and typographical errors of the known locations obtaining an accuracy of 96%. The second module extracted binary and ternary relations between entities to obtain relevant traffic information with a final 87%.

6 Aims, Purposes and Validity

The main objective of almost all the works performed in the literature is obtaining traffic related data from different Social Media. Thus they hope to generate a new constant source of information coming from them. Several works use this information as a complement to that obtained from traffic sensors already installed on the road network. They are considering its users as new sensors to increase the system's range. We also find works in which the extracted information is only used as a support for the existing systems. Finally a few of them only want to show the relations between traffic intensity and social network activity.

WANICHAYAPONG ET AL. [2011], classified traffic information into two types: point and link. First, point information associates with only one point (e.g. a car crash at a crossroad), while link information associates with a road start point and an end point (e.g. a traffic jam between two squares). They want to capture and extract traffic information from all accounts from social networks. The information should be complete enough in order to be useful; hence it should state two important pieces of information of the traffic events: "what" and "where". The extracted traffic events should be informative, factual, and definitive. The last step in their methodology is adding geolocation (latitude and longitude) of start point attribute and end point attribute in the text messages. Geolocation will help visualizing the traffic information on the map. They extracted traffic information from Twitter and classified it. Then, they broadcasted that useful information to help many people plan their routes, avoiding traffic congestion in real-time. This effective classification algorithm can be applied to traffic information from other social networks, not only for Twitter. The study was held in Thailand where more than a Million of tweets were collected for training and testing and even today it is under continuous operation.

D'ANDREA ET AL. [2015] present a real-time monitoring system for traffic event detection from Twitter stream analysis. The system fetches tweets from Twitter, processes them and finally performs the classification of tweets. Their aim is to assign the appropriate class label to each tweet, as related to a traffic event or not. Their proposed system is able to classify SUMs as related to a road traffic event or not and also able to discriminate if traffic is caused by an external event (e.g., a football match, a concert, a political demonstration) or not (e.g. traffic due to congestion or crash). The developed system was installed and tested for the real time monitoring of several areas of the Italian road network, by means of the analysis of the Twitter stream coming from those areas. The aim was to perform a continuous monitoring of frequently busy roads and highways in order to detect possible traffic events in real-time or even in advance with respect to the traditional news media. The system is implemented as a service of a wider service-oriented platform to be developed in the context of the SMARTY project. The service can be called by each user of the platform, who desires to know the traffic conditions in a certain area. The traffic detection system was employed for real-time monitoring (with really good results) of several areas of the Italian road network, allowing for detection of traffic events almost in real time, often before online traffic news web sites. As future work they should integrate their system with an application for analyzing the official traffic news web sites, so as to capture traffic condition notifications in real-time. Thus, their system will be able to signal traffic-related events in the worst case at the same time of the notifications on the web sites.

Existing work on near-term traffic prediction –those that considered forecasting horizons in the range of 5 minutes to 1 hour– relies on the past and current traffic conditions. However, once the forecasting horizon is beyond 1 hour, common in longer-term traffic prediction, these techniques do not work well since additional factors, other than the past and current traffic conditions, start to play important roles. Accurate prediction of both near-term and longer-term traffic conditions can be really useful for those traffic management agencies in order to generate proactive strategies to minimize the effects of traffic congestions. It can also help road users to better plan their trips by avoiding road segments expected to be congested. For the first time, HE ET AL. [2011] examine whether it is possible to use information in online social media to improve longer-term traffic prediction. To this end, they first analyze the correlation between traffic volume and tweet counts with various granularities. Then they propose an optimization framework to extract traffic indicators based on tweet semantics using a transformation matrix, and incorporate them into traffic prediction via linear regression. To accommodate the correlation analysis, they need two data sets: one containing traffic measurements, and the other containing tweet information. They generate the traffic data set by collecting measurements from 943 loop detectors covering the San Francisco Bay area between August 3, 2011 and September 30, 2011 and they also collected tweet data for the same area during the same time period. This results in a total number of 212,145 tweets from 19,435 distinct users. Finally, experimental results using this data demonstrate the effectiveness of their proposed framework.

MEGALLY [2012] develops a system that extracts temporary traffic information from Tweets to be integrated with route planning. The system works in an online manner and processes each tweet once it arrives. She creates and evaluates a supervised machine learning classifier to classify tweets into either tweets related to traffic or tweets not related to traffic. Her system also classifies the tweets into either positive -if it mentions that lanes are clear or road is open- or negative if it mentions that a lane or road is blocked. Then, from traffic related tweets, the location and the nature of the incident (road closed or open) are extracted. Moreover any tweet not related to traffic is discarded from any further processing. Once the location of the traffic perturbation is extracted, the coordinates are used to locate the places where traffic incident occurs and avoid them when providing a route. Probing deeper, the results in her thesis provide a good foundation for future work. A challenging task would be to build the dictionary in an automated manner instead of our manual approach. This will allow the information extraction module to be ported to other areas and other languages. Also, it would be interesting to obtain larger data sets to train and test the classifiers that might lead to better precision and recall values and more generalized tasks.

The system also can be extended to work on a cluster for a better real-time performance. Finally, the system was configured to work for the San Francisco Bay Area.

The main objective of ENDARNOTO ET AL. [2011] was to help people in Jakarta, Indonesia, to get the news of traffic from a reliable source with a very nice presentation. They achieved it by developing a system that can extract the information of traffic from Twitter account of TMC Polda Metro Jaya (police unit in Jakarta) to be presented in a map view by using Google Map and implement it in Android-based mobile application. The tasks were basically to analyze a tweet, get certain information that they need regarding the traffic, and use those information for Android mobile application which will display the traffic condition in map form, with 3 different colors for different traffic conditions (green for normal, yellow for crowded, red for jammed). They chose Android because of the increasing popularity of Android-based mobile phone in Indonesia. The result of their project showed that the information extraction system worked well to extract traffic information from Twitter, and the Android-based mobile application worked well also to display the information in map view. Their experiment has also shown that there is a chance the system will not read the input tweet. So the dictionary, both for vocabulary and rules needs to be updated regularly, especially for the database of place and its coordinate. Future changes need to be done, for example integrate the whole system with the profiling based mobile advertisement, new feature of the Android-based mobile application and new source of traffic information such as Twitter of National Traffic Management Center which means larger scope of information, not only the city of Jakarta but also other area of Indonesia as well.

ACHING ET AL. [2014] propose a strategy to use messages posted in a blogging platform for real-time sensing of traffic related information. Specifically, they use the data that appear in a blog, in Portuguese language, which is managed by a Brazilian daily newspaper on its online edition. They propose a framework based on two modules to infer the location and traffic condition from unstructured, non georeferenced short post in Portuguese. Specifically, the information they deal with is provided by the blog Trânsito, which is managed by the daily newspaper O Estadao de Sao Paulo (OESP), online edition. This service provides short posts in Portuguese, containing traffic information of the city of Sao Paulo. This info comes from users of the transportation system (via Twitter), as well as from manual gathering from various websites (this task is performed by journalists). Their goal is to use the information contained in these short posts to infer traffic condition on several locations. They remark that these posts were not georeferenced, thus the extraction of location information as well as its geolocation was also difficult. Their proposed framework can offer a valid solution to the problem of extracting traffic information from unstructured short posts.

NI [2013] aims to use social media information to assist traffic flow prediction under special event conditions. Specially, a short-term traffic flow prediction model, incorporated with tweet features, is developed to forecast the incoming traffic flow prior sport game events. The motivation of her research is based on the relationship between social media and traffic flow to provide effective models and indicators for traffic prediction in the age that has the largely increased users in social media and vast user-generated contents. By understanding the potential linkage with social media and freeway traffic flow, future researchers can probably identify the social media data as another type sensor for traffic demand, which is updated in real time with no building cost and no need for maintenances. Finally she also argues the efficiency of computation time for social media data processing in practical real-time applications and shows the benefit gained by including social media information in the prediction model and its computational efficiency for potential practical applications.

MEILIN ET AL. [2011] present an application for Traffic Events Detection and Summary (TEDS) based on mining representative terms from the tweets posted when anomalies occur. Their proposed ensemble application contains an efficient TEDS search engine with multiple indexing, ranking, and scoring schemes. This application could benefit both drivers and transportation authorities. Users can search transportation status and analyze traffic events in specific locations of interest. This application allows users to search for traffic information near a specific target location and summarize new traffic events in the vicinity. They have developed a novel application for traffic data, namely wavelet time domain analytics. Once a location is specified, their approach ranks and extracts tweets' top words around the location center according to signal weights and cross correlations in the local neighborhood. Utilizing the proposed signal processing technology, they demonstrate the system's effectiveness by examining traffic and metro travel in the Washington D.C. area. As the collaboration between a citizen's life and social media becomes ever greater, this could have a significant impact on the prediction of traffic flow, travel selection, and other city computing functions. Even all these advantages, the efficiency of their proposed application has not been proved.

ELSAFOURY [2013] proposes a system to use traffic information shared by Twitter messages in a real time manner. It also follows the traffic information sent by @TfLTrafficNews which is an official Twitter account used for reporting about traffic in London. The result of the system is a map showing a highlighted route. This route is the location (road) mentioned in the tweet. The highlight colour depends on the traffic status which is also mentioned in the tweet. The results were tested by comparing it against Google Maps traffic feature. A prototype of the proposed system was implemented in London. To test the system results, the system processed different tweets during different time of the same day. At the same time, screen shots were captured for the same locations on Google Maps traffic results for the same times of issuing and processing the tweets. Comparing both results we can conclude that her research has fulfilled its main objectives.

CARVALHO ET AL. present an initial attempt to use micro-blogging messages posted on Twitter (by users in transit) to perform real-time sensing of traffic-related information. They propose a text classification approach to the problem to automatically identify traffic-related messages posted on Twitter, among the millions of unrelated messages posted by users. Given that the fraction of relevant traffic messages on Twitter is extremely low (< 0,05%), their main challenges involved at this stage are creating a suitable training set for setting up the classifier, and driving the classifier to a reasonable level of precision in identifying relevant messages. All the experiments were performed over a data set of approximately 565,000 Twitter messages crawled from the Portuguese Twitosphere between March and April of 2010. Most of the messages were written in Portuguese but some contain excerpts of other languages, such as English. Results show that, despite the highly-unbalanced example distribution, they were able to almost double the performance of the classifier from the first iteration to the second, and that identifying relevant traffic-related Twitter messages can be achieved, with little human effort involved in creating a training set. Their experiments show that exploring traffic-related information from this media can become an interesting option for complementing information gathered by more traditional sensors (which tend to be limited by several factors such as cost, mobility, availability, etc.). Their future goal is going to be improving classification performance even further. More specifically, they should check how much improvement it is possible to achieve by executing additional iterations of the bootstrapping process.

TOSTES ET AL. presented a system for real-time information about the traffic conditions by analyzing two social sensing sources, one derived from Foursquare, and another derived from Instagram. They try to verify if check-ins can be used as a hint of traffic conditions changes or current situation. They finally found a correlation between traffic congestion and check-ins intensity and showed it in a colored intensity map of Manhattan. More in detail, they have investigated whether Foursquare and Instagram check-ins can be used to signalize congested traffic flow. Through a methodology of five steps, they have showed that these information are correlated. Based on the temporal and the spatial analyses associated with the distribution, they have finally shown that the distribution of check-ins is equal to the distribution of congested traffic flow with a discrepancy error that can be easily calculated, as they have demonstrated. The main problem of this approach is the variation of the threshold's time (time between the congestion and the time that the person arrives at the check-in point). However and as they showed in the paper this time can be easily calculated without being a major problem.

SCHULZ ET AL. [2013] present a solution for a real-time identification of small scale incidents using blogs, thereby allowing increasing the situational awareness by harvesting additional information about incidents. An evaluation based shows that their solution enables the identification of small scale incidents with a high accuracy as well as the detection of all incidents published in real-time Linked Open Government Data. As the average number of tweets identified by their system for each car accident was around ten (with a minimum of only three tweets for one accident), this shows that this approach is capable of detecting incidents with only very few social media posts. In the future they should refine their approach, for example by using more sophisticated NLP techniques, exploring the capability of his approach to detect other types of events, as well as including larger sets of open government data in the evaluation.

As shown above, several systems were finally successfully applied to traffic management in combination with other systems. Moreover their accuracy was also higher –barring a few exceptions– when classifying and extracting traffic related information. For these reasons we can assert both the great validity and use of the information extracted from Social Media using such approaches and techniques.

7 Proposal

Today, people generate enormous amounts of information and mainly through the use of Social Networks. The use of such websites has been widespread and popularised thanks to smartphones and tablets allow their use anywhere, just only needing an Internet connection. In addition, the use of Social Networks is another activity more during the day-to-day of a lot of people, and their use has progressively increased during the last years. For this reason, people are constantly generating new information, in many cases personal, and available for everyone.

On the other hand and as previously mentioned, a new use of Social Networks has recently arisen, or better said, they are seen as a quick and effective new source of information in order to obtain information in real time. As a consequence of their characteristics that allow, only in a few seconds, both write and read a message, real life events may be broadcasted practically live. In addition, today, when big scale occurrences take place, that information related to them has a higher velocity and dynamism through Social Networks as opposed to traditional sources of information. However, not all are advantages because, as a result of the enormous number of users giving information about such event, the veracity of a large percentage of the information is not corroborated, raising doubts and lack of credibility. However, the great potential of such tools is starting to be visible, due to applying a correction or weighting coefficient on the sample, considered as such all the information related, it is undeniable that for big scale events may be extracted really useful information from these sources, representing a huge step forward.

However, the problem arises when such information is based on a minor scale event, either small or medium scale, with a number of SUMs far lower than the previous case, and thus with a database also limited. As a consequence, and as seen before, once applied the weighting coefficient on the sample, the latter could be really small and then without providing added value or little information about the event. Also should be mentioned that those weighting coefficient used for big and small or mid-scale events would not be the same, due to users are more interested in big-scale events and therefore the uncertainty of the sample is also bigger. Nevertheless, what is clear is that for small scale events the main problem is that, probably, such generated information would be not enough to efficiently detect and describe its origins and features.

On the previously defined scenarios is where this paper is focused and based. As is well known, if this information is processed in an efficient way, very useful information could be obtained from such processing. However, and connecting with the topic of interest, those events related with traffic (congestion, accidents, traffic conditions) are usually small or mid-scale events instead of big-scale. In case of, for example, a massive or multiple traffic accident or long and chaotic congestion, these events could be truly considered as big scale.

As opposed to the latter, the true utility of the different information extraction methods mentioned in this paper are based on small or mid-scale events. For this reason, such methods are early limited by if the coverage of a traffic related events is high enough to allow data extraction modules to obtain valuable information.

On the other hand, if the event takes place in a punctual and easy point to identify (such as a bridge, street or intersection) and the levels of generated SUMs in Social Networks are reasonable, there are powerful data processing tools to extract information from these messages. As mentioned and explained in previous sections of this paper, exist multiple methods and tools to process information present in Social Networks, and once the processing has been applied to the sample, extract really valuable information regarding a traffic related event. Besides, and in order to increase the percentage of certainty, multiple methods or algorithms could be used to obtain better and reliable results. The main problem is the necessity of a substantial volume of information regarding the covered event, and on the other hand the necessity to know how to configure such systems and tools in order to extract information. Once again, the purpose of such paper is, based on the literature, the identification of the main methods used on it. For this reason, knowing these methods is only the first step, due to they need to be trained, in some cases supervised and even based on specific dictionaries to correctly identify valuable information. This is intended to emphasise that, the knowledge bases not only on those methods used to such purpose, but also in having a wide know-how or expertise on this field. Consequently, the efficiency or savings on resources could not be so, perhaps, due to such informatics procedures require additional human capital and computer equipment, increasing costs.

On the other hand, once such procedures are installed and configured, their reliability is in general very effective. Due to several works mentioned in the literature had, later, a practical application in real traffic, they allowed to demonstrate their high percentages of reliability. However, in the majority of cases they were used as a supplementary tool of those already existing, increasing the range of the whole system and the available information. In addition, and as has already seen, none of the studied cases have used –individually– the proposed procedure to detect traffic anomalies or to extract traffic related information. But it is also necessary to mention that, such procedures cannot be used individually due to already exist, in general, systems to monitor and control traffic conditions, which allow obtaining traffic information. However, an interesting question would be if these procedures could be used as an autonomous unit, independent from any other, as a new method in future streets and highways. Another important question could be if they could be enough useful and profitable to implement them as new ways of obtaining information from traffic.

In order to answer all these questions, added to it is necessary to summarize both the main advantages and disadvantages of such proposed methods, it has been deemed relevant to undertake a SWOT analysis on them. Consequently, the next section bases on such analysis with the aim of clarifying doubt and answering all pending questions.

7.1 SWOT Analysis

Due to we want to answer the questions discussed in the previous section through reasoned arguments, this section is based on, firstly a detailed analysis of the main advantages and disadvantages of using such information methods, followed by the answers to such questions from the previous analysis. More formally, such questions could also arise to those companies or entities, concerned with how these new information extraction techniques could improve their day-to-day, and therefore they are considering this new project. Their decision would be then, if positive, implementing these procedures to increase their traffic related data (with their obvious costs), or if negative, continuing only with traditional systems or sensors to obtain traffic information. In this case, we have proposed another scenario, but this one could be perfectly truthful and applicable. For this reason, the company or entity (from now on company) should take a decision between if implementing the project or not, according to which situation is more favorable to it. Consequently, the company should analise in detail both alternatives, having in mind which one will have a higher expected return. Due to decision making is, in some companies, used to be based on non-rational opinions without judgment, such organizations have to base this decision in well structured processes in order to increase the knowledge about the alternatives, and consequently decrease the negative effect of these decisions. This is the root of the importance of SWOT analysis, a procedure widely used when it comes about knowing the current situation of a project or company, in order to plan future strategies. Therefore, the objective of the following SWOT analysis is solving if all these procedures, used to extract traffic information from Social Media, are so useful and reliable for those companies wishing to implement them as a part of their information systems.

The previously mentioned SWOT analysis (acronym for Strengths, Weaknesses, Opportunities and Threats) is based on the study of the current situation of a company or project, such as is the case, analysing, on the one hand its internal characteristics (both strengths and weaknesses) and on the other hand its characteristics but from an external standpoint (opportunities and threats). The result of this analysis is usually presented in form of a square matrix, which allows analysing and comparing its different components in a quick and visual manner.

Now the first step is to determine, according to the analysis, the internal characteristics of the project and consequently its strengths and weaknesses. On the one hand, its strengths are those qualities of the project that allow the company to gain a competitive advantage over the rest of projects, or as in this case, the rest of alternatives. In this case, the only possible alternative is that which does not apply these systems, and therefore the company continues its operating activity with the same traffic information systems. On the other hand, weaknesses are referred to all those aspects that result in a competitive disadvantage or lack of competitiveness compared to other projects.

Once these first two concepts have been explained, the following step is their determination. The main strengths of such project are, respectively, it allows for increasing available information about traffic, is also able to foresee –in some cases– traffic occurrences and the lifespan of such systems is much longer than traditional methods. It is clear that available information or coverage, once these systems will be implemented, is going to be bigger than before. This is the result of Social Networks and their users, which through them provide traffic related information that can be used as complement for those coming from traditional sources. An example of this is that such SUMs may provide information about waiting times in traffic jams, traffic conditions, number of vehicles involved in accidents and even condition of injured persons and existence of deaths. In addition, SUMs related to traffic incidences, as well as volume rises in published SUMs in a particular place, may be used as traffic forecast. Consequently, if such messages are correctly analised, once this rise has been detected according to a given pattern, they could act accordingly in order to minimise or even remove future accidents. Therefore the role of these methods in not only limited to traffic information extraction, but also interpreting it to act accordingly. With regard to the last point, such systems may be used as long-term systems, as opposed to traditional systems that are restricted by smaller lifespan of their sensors and electronic devices. Due to the characteristic design of those new methods used to extract information from SUMs, their only maintenances would be their regular upgrading, in order to keep them up-to-date considering new SUM patterns or reviewing used dictionaries to extract relevant information.

On the other hand, the main weaknesses of this methodology are that, in many cases the available SUMs will be very little, as well as the information provided by such messages, being mostly considered as noise. Besides, additional resources should be employed in order to implement these procedures, which require special training and additional systems or methods, increasing both human –to supervise and train– and computer resources –to install and operate the system–. Lastly, operation and maintenance costs could be even higher than in traditional methods due to it requires specific and trained personnel. Besides, there are additional devices that need to be monitoring and supervised. As previously mentioned, traffic related occurrence are used to be, mostly, small scale events. Consequently, in case of accidents that do not involve several vehicles (thus considered as big scale) and usual traffic incidents as traffic jams, all those SUMs published and used to feed the system could be not enough to obtain valuable information. In addition such messages could be mostly considered as noise, due to such events are common and repetitive and the users may progressively ignore relevant information in their messages (in favor of moods). On the other hand, it is said that these extraction systems would allow reducing maintenance costs and increase lifespan as opposed to traditional systems. Whilst the latter may be considered true in the long term, what has not been considered yet is the effect of their implementation in the company and its financial statements. This means that, if the company wants to carry through with the project, additional provisions need to be done in order to install and train the system (human resources) and support such new programs (computer resources). Thus, this need of additional resources implies, in the short term, a higher capital expenditure than expected –without considering the project– and small long term savings. Therefore this could be seen as a major obstacle by many companies. Finally, as in the strengths, maintenance costs are also considered due to it is not clear that they decrease with these new methods. Once again, it is based on that they are going to need specific and trained personnel to supervise them and additional computer devices, increasing costs.

Once finished the analysis of the internal characteristics, the second step is to determine those respective from an external standpoint. This second part of the analysis considers the project from its surroundings, as well as the different opportunities and threats that could appear in it. Consequently, the objective now is determining, on the one hand those facts or agents that could influence the project, and on the other hand which effects could have these parties in the project in terms of easing or complicating to reach its proposed goals. This means that, there are circumstances or facts present in its surroundings that, sometimes, may represent an opportunity to the project. On the contrary, may exist negative scenarios or situations, considered as threats, which could worsen the project.

The first step of this second analysis is identifying all those opportunities that could arise as a consequence of the implementation of this new methods or procedures. One opportunity is that sources do not have any kind of restriction according to temporal and geographical reasons, having continuous flow of information. For this reason, provided that users are publishing information about traffic and, very important, ready to inform about such traffic related event, the system will have a continuous database where to feed. In addition, this source of information has free access through API devices, which are completely independent from Social Networks, and without any cost. As previously mentioned, these applications are not only used to extract information from Twitter but also from any kind of Social Network. Finally, it needs to be underlined that the use of such networks has been widespread and actually it is totally usual in the society as a whole. Therefore, these Social Networks seen as sources of information have a considerable lifespan in the long term.

Lastly, it only remains the identification of the main threats that this project had to deal with. The main threat, as it has been said repeatedly, is the availability of traffic related SUMs that contain relevant and reliable information about an event. This basically means that, one the one hand, it may be not enough volume of SUMs to feed the system, and therefore the latter is not able to extract any kind of information from them. On the other hand, even if this volume is significant, or better said, representative to not forget with large, the information contained in the SUM may not provide additional information from those obtained from traditional sources. For this reason, the utility of such methods would be highly limited to incoming volumes of traffic information, as well as its conditions. Moreover, and due to this project is based on new technologies, its future is uncertain. This means that, even it could be in contradiction with one of the opportunities previously mentioned, the track-record of such Social Networks is, indeed, very short. In other words, their expansion during the first years was undeniable, but once such Networks reach higher degrees of maturation (e.g., reduction in the number of new users, widely criticised in case of Twitter), they might be replaced by others or even become obsolete. The purpose of this is to emphasise that, due to most of the proposed methods are based on Twitter, a change in usage patterns and preferences of their users may cause that the system become useless. One example is Foursquare and Instagram, in which you can show the specific position where you are, together with text and image. Taking advantage of this fact, TOSTES ET AL. in their paper about traffic condition analysed from Foursquare and Instagram data, established a correlation between such posts (also called *check-ins*) and traffic condition, allowing them to forecast traffic jams. However, with the passing of years Foursquare began to be become obsolete in contrast to Instagram. In this case, their paper was based on two different Social Networks, but if this had not been the case, the system and all resources employed to design it would have also become wasted, once again as a consequence of such changes in usage patterns and preferences of their users.

Consequently, and once identified and defined the different components of SWOT analysis, one should wonder if the previous questions, about the feasibility of the project for those parties wishing to carry it out, have been answered. Therefore, in order to answer such questions through reasoned argument, it is necessary to distinguish between two possible scenarios, otherwise they would lack of content. Establishing these two proposals we want to differentiate between those companies that manage road infrastructures but, on the one hand they are very long, such as a highway, or, on the other hand those that are focused on specific sections, such as a bridge or a tunnel. This means that, in the first case they are going to manage low to medium traffic incidents, without constant traffic jams or accidents. One example might be those highways with fluid flows of traffic and low concentration of accident blackspots. For this first scenario, it is clear that the best decision is not carrying out the project due to their current information systems provide a relevant degree of information. Consequently, the cost of adding such procedures to the system would not revert increasing the information of traffic, in other words, such resources would provide almost the same information as the currents systems, due to the characteristics of the infrastructure and its low volumes in traffic related SUMs. Thus, neither it cost nor it application would be balanced with the rise in traffic information.

The second scenario considers, for its part, those road infrastructures that, given their high number of daily traffic incidents, which are also characterized by their occurrence or repetition in easily identifiable points (e.g., street, bridge, tunnels or even an exit of a highway), and could involve multiple users ready to publish SUMs containing traffic information. In this case and differently from the previous, the incoming flow of SUMs would be much greater also because of the concentration of users. Besides, and due to they happen in know locations, additional information to that provided by traditional sources could be extracted from SUMs. This information, if correctly analysed, might also forecast traffic flows helping to manage future incidents and consequently decrease the number of accidents and traffic jams. However, those companies wishing to implement such information extraction systems have to be aware that they entail an economic impact. However, and as seen in the literature, is in these cases when such projects are truly useful and their implementation is highly recommended. Once again and as said before, this project implies a cost, and therefore an in-depth study should be performed in order to assess its cost-effectiveness and to decide its usefulness for the future.

In conclusion, these proposed methods to extract traffic information are really interesting, always, as complements of traditional systems and provided that the scope of application is similar to the last scenario, as above defined. On the contrary, if the project is focused on a location with low incoming flows of SUMs, its implementation is not advisable.

8 Conclusion

The use of Social Media sites has become widespread and popular during recent years, turning into a channel for real-time information. People intensely use social networks to report (personal or public) real-life events happening around them. Recently, social networks and media platforms have also been used as a source of instant information for the detection of real-world occurrence that happens in a specific time and space. However, event detection from social networks analysis is a more challenging problem than event detection from traditional verified and well formatted media. In addition, traffic related events are small-scale events and consequently have a small number of SUMs related to them.

Nevertheless, nowadays there are different methods and techniques to extract traffic related information from Social Media. There are some differences between the different methods, but we can distinguish between two main methods: Named Entity Recognition and Part of Speech. The fist one requires training methods to classify the incoming information, lengthening the process but it obtains great results. The second one, based on a simpler analytical basis, also reach good results and being competent regarding to data extraction. We have seen the different and high accuracies obtained by the different works in the literature. Also the use of correlations between Social Media activity and traffic intensity has answered the question that it is possible to predict traffic patterns. So not only we can extract valuable information about traffic related events but also can predict (with a certain margin of error) future traffic anomalies.

This paper has been written to summarize the different techniques used in the literature and to show the accuracies obtained processing such data. It is noteworthy that, due to the increase in number of Social Media users, we can have access to traffic related information even sooner than the one coming from the sensors. Thus, an accurate use of these extraction techniques can generate great benefits to the agency responsible for its operation. Do not forget that today we are living in the Data era and a proper treatment of such information represents a great value.

Finally, and as a result of the SWOT analysis, the main advantages and disadvantages of these procedures (from different standpoints) are brought to the forefront in order to evaluate in which scenarios they could be effectively used. The conclusion is that these proposed methods are really interesting to extract traffic information, always, as complements of traditional systems and provide that there were enough flows of information published in Social Networks. On the contrary, if the project is focused on location with low incoming flows of SUMs, its implementation is not advisable.

List of References

ATEFEH F.; KHREICH W. [2015]: A survey of techniques for event detection in Twitter," Comput. Intell., vol. 31, no. 1.

P. RUCHI AND K. KAMALAKAR [2013]: ET: Events from tweets,. 22nd Int. Conf. World Wide Web Comput., Rio de Janeiro, Brazil.

T. SAKAKI, M. OKAZAKI, Y.MATSUO [2013]: Tweet analysis for real-time event detection and earthquake reporting system development. IEEE Trans. Knowl. Data Eng., vol. 25, no. 4, Apr. 2013.

T. SAKAKI, Y. MATSUO, T. YANAGIHARA, N. P. CHANDRASIRI, K. NAWA [2012]: Real-time event extraction for driving information from social sensors. IEEE Int. Conf. CYBER, Bangkok, Thailand, 2012.

N. WANICHAYAPONG, W. PRUTHIPUNYASKUL, W. PATTARA-ATIKOM, P. CHAOVALIT [2011]: Social-based traffic information extraction and classification. 11th Int. Conf. ITST, St. Petersburg, Russia, 2011.

J. HURLOCK M. L. WILSON [2011]: Searching twitter: Separating the tweet from the chaff. 5th AAAI ICWSM, Barcelona, Spain, 2011.

S. WEISS, N. INDURKHYA, T. ZHANG, F. DAMERAU [2004]: Text Mining: Predictive Methods for Analyzing Unstructured Information. Berlin, Germany: Springer-Verlag, 2004.

A. HOTHO, A. NÜRNBERGER, G. PAAß [2005]: A brief survey of text mining. LDV Forum-GLDV J. Comput. Linguistics Lang. Technol., vol. 20, no. 1, May 2005.

M. W. BERRY; M. CASTELLANOS [2004]: Survey of Text Mining. New York, USA: Springer-Verlag, 2004.

H. TAKEMURA; K. TAJIMA [2012]: Tweet classification based on their lifetime duration. 21st ACM Int. CIKM, Shanghai, China, 2012.

A. SCHULZ, P. RISTOSKI, H. PAULHEIM [2013]: I see a car crash: Real-time detection of small scale incidents in microblogs. ESWC 2013 Satellite Events, vol. 7955. Berlin, Germany: Springer- Verlag, 2013.

M. KRSTAJIC, C. ROHRDANTZ, M. HUND, A. WEILER [2012]: Getting there first: Real-time detection of real-world incidents on Twitter. 2nd IEEE Work Interactive Vis. Text Anal.—Task-Driven Anal. Soc. Media IEEE VisWeek, Seattle, WA, USA, 2012.

C. CHEW AND G. EYSENBACH [2010]: Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak.PLoS ONE, vol. 5, no. 11, Nov. 2010.

B. DE LONGUEVILLE, R. S. SMITH, G. LURASCHI [2009]: OMG, from here, I can see the flames!: A use case of mining location based social networks to acquire spatio-temporal data on forest fires.. Int. Work. LBSN, 2009 Seattle, WA, USA.

J. YIN, A. LAMPERT, M. CAMERON, B. ROBINSON, R. POWER [2012]: Using social media to enhance emergency situation awareness. IEEE Intell. Syst., vol. 27, no. 6, Nov./Dec. 2012.

P. AGARWAL, R. VAITHIYANATHAN, S. SHARMA, G. SHRO [2012]: Catching the long-tail: Extracting local news events from Twitter. 6th AAAI ICWSM, Dublin, Ireland, Jun. 2012.

F. ABEL, C. HAUFF, G. J. HOUBEN, R. STRONKMAN, K. TAO [2012]: Twitcident: fighting fire with information from social web streams. ACM 21st Int. Conf. Comp. WWW, Lyon, France, 2012.

R. LI, K. H. LEI, R. KHADIWALA, K. C.-C. CHANG [2012]: TEDAS: A Twitterbased event detection and analysis system. 28th IEEE ICDE, Washington, DC, USA, 2012.

Y. ZHOU; Z. W. CAO [2011]: Research on the construction and filter method of stop-word list in text preprocessing. 4th ICICTA, Shenzhen, China, 2011, vol. 1.

G. SALTON; C. BUCKLEY [1998]: Term-weighting approaches in automatic text retrieval. Inf. Process. Manage., vol. 24, no. 5, Aug. 1988.

L. M. AIELLO [2013]: Sensing trending topics in Twitter. IEEE Trans. Multimedia, vol. 15, no. 6, Oct. 2013.

C. SHANG, M. LI, S. FENG, Q. JIANG, J. FAN [2013]: Feature selection via maximizing global information gain for text classification. Knowl.-Based Syst., vol. 54, Dec. 2013.

Y. APHINYANAPHONGS [2014]: A comprehensive empirical comparison of modern supervised classification and feature selection methods for text categorization. J. Assoc. Inf. Sci. Technol., vol. 65, no. 10, Oct. 2014.

J. PLATT [1999]: Fast training of support vector machines using sequential minimal optimization. In Advances in Kernel Methods: Support Vector Learning, B. Schoelkopf, C. J. C. Burges and A. J. Smola, Eds. Cambridge, MA, USA, MIT Press, 1999.

G. H. JOHN, P. LANGLEY [1995]: Estimating continuous distributions in Bayesian classifiers. 11th Conference Uncertainty Artificial Intelligence, San Mateo, CA, 1995.

J. R. QUINLAN [1993]: C4.5: Programs for Machine Learning. San Mateo, CA, USA: Morgan Kaufmann, 1993.

D. W. AHA, D. KIBLER, M. K. ALBERT [1991]: Instance-based learning algorithms. Mach. Learn., vol. 6, no. 1, Jan. 1991.

C. CORTES, V. VAPNIK [1995]: Support-vector networks. Mach. Learn., vol. 20, no. 3, Sep. 1995.

T. T. COVER ,P. E. HART [1967]: Nearest neighbour pattern classification. IEEE Trans. Inf. Theory, vol. IT-13, no. 1, Jan. 1967.

E. D'ANDREA, P. DUCANGE, B. LAZZERINI F. MARCELLONI [2015]: Real-Time Detection of Traffic From Twitter Stream Analysis. February 10, 2015.

S. VERMA, S. VIEWEG, W. CORVEY, L. PALEN, J. MARTIN, M. PALMER, A. SCHRAM [2011]: Natural language processing to the rescue? extracting "situational awareness" tweets during mass emergency. 5th International AAAI Conference on Weblogs and Social Media (ICWSM), Barcelona, 2011.

E. BENSON, A. HAGHIGHI, BARZILAY [2011]: Event discovery in social media feeds. 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Vol. 1. Association for Computational Linguistics, Stroudsburg, 2011.

B. LOCKE, J. MARTIN [2009]: Named entity recognition: adapting to microblogging. Senior Thesis, University of Colorado, Colorado, 2009.

A. RITTER, S. CLARK, MAUSAM, O. ETZIONI [2011]: Named entity recognition in tweets: an experimental study. 2011 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Edinburgh, Scotland, 2011.

J. HE, W. SHEN, P. DIVAKARUNI, L. WYNTERZ,R. LAWRENCE [2011]: Improving Traffic Prediction with Tweet Semantics. Computer Science Department, Stevens Institute of Technology, New York, 2011.

M. MEGALLY [2012]: Information Extraction from Social Media for Route Planning. Master's thesis Nr. 3413, Universität Stuttgart, November 15, 2012.

S. K. ENDARNOTO, S. PRADIPTA, A. S. NUGROHO, J. PURNAMA [2011]: Traffic Condition Information Extraction & Visualization from Social Media Twitter for Android Mobile Application. Faculty of Information Technology, Swiss German University, Tangerang, Indonesia, 2011.

J. L. ACHING, T. B. F. DE OLIVEIRA, A. L. C. BAZZAN [2014]: Traffic Information Extraction from a Blogging Platform using Knowledge-based Approaches and Bootstrapping. Instituto de Informática, Universidade Federal do Rio Grande do Sul, Brazil, 2014.

M. NI [2013]: Using Social Media to Predict Traffic Flow under Special Event Conditions. Faculty of the Graduate School of the State University of New York, Buffalo, 2013.

L. MEILIN, F. KAIQUN, L. CHANG-TIEN, C. GUANGSHENG, W. HUIQIANG [2011]: A Search and Summary Application for Traffic Events Detection Based on Twitter Data. China, 2011.

B. PAN, Y. ZHENG, D.WILKIE ,C. SHAHABI [2013]: Crowd Sensing of Traffic Anomalies based on Human Mobility and Social Media. 2013.

F. A. ELSAFOURY [2013]: Monitoring urban traffic status using Twitter messages. Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente, the Netherlands, February, 2013.

S. CARVALHO, L. S. ROSALDO J. F. ROSSETTI: Real-Time Sensing of Traffic Information in Twitter Messages", Portugal.

A. I. J. TOSTES, T. H. SILVA, F. DUARTE-FIGUEIREDO: Studying Traffic Conditions by Analyzing Foursquare and Instagram Data.

D. MCHUGH [2015]: Traffic Prediction and Analysis using a Big Data and Visualisation Approach. Department of Computer Science, Institute of Technology Blanchardstown, March 10, 2015.

JEFFREY STRICKLAND [2015]: Predictive Analytics using R. Lulu.com

PART OF SPEECH: available at: https://en.wikipedia.org/wiki/Part_of_speech

List of Abbreviations

SUM	Status Update Message
ITS	Intelligent Transportation System
NLP	Natural Language Processing
CDE	Crime and Disaster-related Event
SVM	Support Vector Machine
NB	Naïve Bayes
k-NN	k-Nearest Neighbors
Maxent	Maximum Entropy
NER	Named Entity Recognition
PoS	Part of Speech
API	Application Programming Interface

List of Figures

List of Tables

Declaration concerning the Bachelor’s Thesis

I hereby confirm that the presented thesis work has been done independently and using only the sources and resources as are listed. This thesis has not previously been submitted elsewhere for purposes of assessment.

Munich, September 24th, 2015

Diego Pacho Toubes